



FACULTAD DE ESTADÍSTICA

MÁSTER DE MINERÍA DE DATOS
E INTELIGENCIA DE NEGOCIO

**PERFILES DE INNOVACION EN LAS EMPRESAS
ESPAÑOLAS EN LOS PERIODOS 2005-2007 y 2009-2011**

Autor: Mónica Pérez Jiménez
Tutores: Javier Amores Salvadó y Gregorio Martín de Castro

Curso Académico 2014/2015

13 de noviembre de 2015

1.	Introducción	4
1.1.	Innovación y éxito innovador	4
1.1.1.	¿Qué es innovación?.....	4
1.1.2.	Clasificación.....	5
1.1.3.	Etapas	7
1.1.4.	Los emprendedores y las fuerzas motrices de la innovación	7
1.1.5.	El proceso de innovación	8
1.1.6.	Medición de la innovación	8
1.1.7.	Innovación en servicios	8
1.1.8.	Innovación y ventaja competitiva	10
1.1.9.	Factores que llevan a la innovación	10
1.1.10.	Factores que influyen en el logro del éxito innovador	11
1.2.	Base de Datos	14
2.	Objetivos.....	15
3.	Metodología Empleada.....	15
3.1.	Herramientas.....	15
3.2.	Metodología.....	16
3.2.1	Proceso ETL.....	16
3.2.2.	Segmentación	17
4.	Desarrollo del Trabajo	18
4.1.	Comprensión de los datos.....	18
4.2.	Selección de fechas de referencia.....	20
4.3.	Creación de variables sintéticas preliminares.....	21
4.3.1.	Tipo de empresa	21
4.3.2.	CNAE.....	21
4.3.3.	Concentración de la industria.....	22

4.3.4.	Categorización de madurez	22
4.3.4.	Ratio de innovación y volumen de negocio	23
4.3.5.	Variaciones o Deltas.....	23
4.3.6.	Otras variables.....	23
4.4.	Definición del Público Objetivo	24
4.5.	Creación de variable objetivo	24
4.6.	Segmentación de empresas 2005-2007	25
4.6.1.	Análisis no jerárquico.....	27
4.6.2.	Árbol de clasificación.....	28
4.6.3.	Resultados	32
4.7.	Segmentación de empresas 2009-2011	33
4.7.1.	Análisis no jerárquico.....	33
4.7.2.	Árbol de clasificación.....	34
4.7.3.	Resultados	35
4.8.	Comparativa	37
5.	Conclusiones.....	39
6.	Próximos Pasos y Principales Limitaciones	43
7.	Bibliografía	45
Anexo I – CNAEs y equivalencias 2005-2007.....		48
Anexo II – CNAEs y equivalencias 2009-2011		50
Anexo III – Índice concentración 2005-2007.....		52
Anexo IV – Índice concentración 2009-2011.....		54
Anexo V – Árbol de Decisión 2005-2007		56
Anexo VI – Árbol de Decisión 2009-2011.....		57
Anexo VII – <i>Scripts</i> SQL Server		58

1. Introducción

Con este trabajo se pretende realizar una segmentación de las empresas españolas según su perfil innovador, con el fin de identificar las características comunes a un proceso innovador de éxito.

Además, se ha querido analizar la influencia de la crisis económica española en los perfiles de innovación de las empresas. Para ello se ha dividido el histórico en dos periodos de análisis: el periodo 2005-2007 y el periodo 2009-2011. Con ello se podrá realizar el estudio de los procesos innovadores antes y durante la crisis económica, así como sus posibles diferencias y analogías.

Con el fin de comprender qué es un proceso innovador, su influencia en los resultados y, en consecuencia, en la estrategia corporativa de la empresa, se comenzará este trabajo a través de un estudio breve del *estado del arte* de la innovación.

1.1. Innovación y éxito innovador

1.1.1. ¿Qué es innovación?

El concepto de innovación fue introducido por J.A. Schumpeter en su libro *The theory of economic development*, publicado en 1934 (Camisón, 2011). Schumpeter veía la economía como un ciclo circular en equilibrio y en competencia perfecta en su estado estacionario (Hagerdoorn, 1996). La innovación actuaba como agente disruptivo que introducía desequilibrio en ese ciclo, permitiendo al innovador o emprendedor obtener beneficios. La ventaja competitiva generada es de carácter temporal, ya que cualquier innovación introducida por un competidor volverá a generar un desequilibrio, haciendo

que el primer innovador pierda su posición de monopolio (Steliana Ciote, 2012). La economía volverá a su equilibrio cuando los efectos de las innovaciones introducidas se atenúen.

El concepto schumpeteriano de innovación no solo se refiere a nuevos productos, si no a “nuevas combinaciones”, es decir, la novedad o mejora de productos, métodos de producción, *inputs* o formas de organización (Hagerdoorn, 1996). Aunque la teoría de innovación de Schumpeter ha sido discutida por autores posteriores, especialmente en su aplicación a servicios, ha sido la base del concepto de innovación como hoy lo conocemos (OECD y Eurostat, 2005). En la actualidad, la innovación es el proceso de transformar un problema u oportunidad en productos o servicios nuevos o mejorados, a través de la generación, aceptación y explotación comercial de nuevas ideas (Hall y Williams, 2008; Parsons y Rose, 2010). Los procesos de innovación no tienen que ser generados por la propia empresa, si no que pueden ser adquiridos a agentes externos (OECD y Eurostat, 2005).

Las empresas, por tanto, llevarán a cabo procesos de innovación con el fin de alcanzar ventaja competitiva o mantener su competitividad como requisito fundamental de supervivencia en un entorno de rápidos cambios y sobreabundancia de información (OECD y Eurostat, 2005).

1.1.2. Clasificación

Existen numerosas formas de clasificar las innovaciones que tienen lugar en la empresa. En primer lugar, se distinguen cuatro grados de innovación según el nivel de novedad que se introduzca, de acuerdo al informe publicado por la OECD y Eurostat en 2005: innovaciones en la empresa, innovaciones a nivel el mercado, innovaciones para toda la sociedad e innovaciones disruptivas.

Otra posible forma de clasificar los procesos innovadores es de acuerdo a su objetivo, como también se detalla en el Manual de Oslo (OECD y Eurostat, 2005). Según su objeto, las innovaciones pueden ser innovaciones de producto, innovaciones de proceso, innovaciones organizacionales e innovaciones en marketing. Otros autores (Hjalager, 2009) distinguen tipos adicionales de innovación, como las innovaciones

institucionales, o realizan clasificaciones alternativas. Es el caso de Parsons y Rose (2010), las innovaciones se clasifican en innovaciones en productos y servicio, innovaciones de proceso, innovaciones radicales o disruptivas e innovaciones incrementales. Ha de tenerse en cuenta que las distinciones entre las diferentes categorías especificadas son subjetivas, y una innovación puede clasificarse más de una categorías. Las innovaciones tecnológicas no se consideran una categoría per se, ya que pueden encontrarse en cualquiera del resto de categorías.

De acuerdo con la clasificación del Manual de Oslo (OECD y Eurostat, 2005):

- *Innovaciones de Productos y Servicios*, son aquellas que implican un cambio significativo en los atributos de un bien o servicio; ya sea la mejora de un bien o producto existente o la creación de uno nuevo (OECD y Eurostat, 2005). Es uno de los tipos de innovación más buscados por las empresas, ya que es fácilmente percibido por su cartera de clientes potenciales (Hjalager, 2009). El éxito de estas innovaciones se miden habitualmente en el incremento de ventas derivadas del producto innovador.
- *Innovaciones de Proceso*, son aquellas que implican un cambio significativo en los métodos usados, ya sea producción o distribución (OECD y Eurostat, 2005) con el propósito de incrementar eficiencia y productividad (Hjalager, 2009).
- *Innovaciones Organizacionales*, o innovaciones directivas (Hjalager, 2009), son aquellos cambios que la empresa implementa en la organización (OECD y Eurostat, 2005), como la reorganización de departamentos internos, la formación de empleados o iniciativas destinadas a incrementar la satisfacción y bienestar de los trabajadores.
- *Innovaciones en Marketing*, se refieren a la implementación de nuevas técnicas de mercadotecnia (OECD y Eurostar, 2005; Hjalager, 2009). Esta categoría ha de ser considerada con precaución, ya que únicamente es correcta cuando lleva consigo un cambio significativo – ya sea en comunicación o en la relación de la firma innovadora con el cliente, entre otros.

- *Innovaciones institucionales*, a pesar de no estar consideradas en el Manual de Oslo (OECD y Eurostat, 2005), son aquellas relacionadas con el diseño e implementación de estructuras colaborativas nuevas o mejoradas, así como cambios en el marco legal (Hjalager, 2009).

1.1.3. Etapas

La clave del impacto económico de los cuatro grados de innovación anteriores es la difusión que tiene lugar en el mercado (OECD, 2005). Schumpeter distinguía cuatro etapas del proceso de innovación: invención, innovación, difusión e imitación (Steliana Ciotte, 2012), y solo cuando estas cuatro etapas tienen lugar el proceso de innovación se traducirá en desarrollo económico (OECD y Eurostat, 2005). La razón de que el proceso de innovación que lleva a cabo una entidad se traduzca en beneficios a la sociedad en su conjunto radica en el carácter social de la innovación: la innovación no solo requiere de la creatividad necesaria para desarrollar una idea exitosa (investigación), sino que es necesaria su comercialización, con lo que la sociedad se beneficia de las mejoras introducidas y del conocimiento generado.

1.1.4. Los emprendedores y las fuerzas motrices de la innovación

Los cambios económicos que tienen lugar en el ciclo de Schumpeter descrito en apartados anteriores, y el desarrollo económico que ello implica, tienen lugar gracias a la figura del emprendedor (Hagerdoom, 1996). De acuerdo a Hjalager (2010), los emprendedores se consideran *destructores creativos*, ya que su capacidad de innovación permite la generación de un cambio radical en las necesidades e intereses del mercado.

Esto tiene lugar gracias al profundo conocimiento que tiene el emprendedor de la empresa, sus clientes y su entorno competitivo, así como de su entorno económico, social e institucional. Gracias a ello, no solo pueden considerar cuidadosamente qué necesita la sociedad, sino que también disponen de las herramientas, conocimientos y experiencia necesarios para transformar una oportunidad en rentabilidad para la empresa. Esta característica hace de ellos un eslabón fundamental en la transformación de creatividad en innovación (Parsons y Rose, 2010).

1.1.5. El proceso de innovación

La innovación no está limitada al departamento de I+D de la empresa, si no que implica a la empresa en su conjunto y ha de estar firmemente radicada en su estrategia corporativa. Es por ello que el proceso de innovación que la empresa lleva a cabo implica actividades tanto científico-tecnológicas, como actividades financieras, comerciales y de organización.

Cada uno de los eslabones del proceso de innovación es necesario para desarrollar e implementar una innovación. Las actividades que se enmarcan en el proceso de innovación pueden clasificarse en tres categorías, de acuerdo a su madurez: innovaciones exitosas, innovaciones en proceso e innovaciones abandonadas o desechadas. Una empresa podrá ser considerada innovadora cuando la innovación haya sido implementada, independientemente de su éxito o no (OECD y Eurostat, 2005).

1.1.6. Medición de la innovación

A día de hoy, la información relativa a innovación es limitada y no se dispone de métodos adecuados para su medición (OECD & Eurostat, 2005; Camisón, 2011; DTI, 2007; Orfila-Sintes & Mattson, 2007), y muy especialmente en el sector servicios, que por su naturaleza es de más difícil medición que el sector industrial. No obstante, en los últimos años la investigación de la innovación ha cobrado una especial importancia. Gracias a ello se han dedicado numerosos esfuerzos a mejorar la comprensión de los factores, causas y etapas, así como la medición de la innovación, como es el caso del Manual de Oslo, publicado en 2005 por la Organización para la Cooperación y el Desarrollo (OECD) y la Oficina Europea de Estadística, Eurostat. A través de la comprensión y análisis del proceso de innovación y de generación, explotación y difusión de conocimiento, será posible incrementar la productividad y beneficios de las empresas, de lo que se beneficiará la sociedad en su conjunto.

1.1.7. Innovación en servicios

A pesar de que a lo largo del desarrollo de este trabajo se tendrán en cuenta empresas de todos los sectores económicos, cabe destacar las peculiaridades del sector servicios, del que por su naturaleza, y por los escasos trabajos de investigación académica

comparado con el resto de sectores, se considera que se podrán extraer conclusiones más interesantes (DTI, 2007; Sundtbo et al., 2007). El volumen de publicaciones de procesos de innovación en servicios es menor, ya que el estudio de la innovación ha estado centrado tradicionalmente en industria (DTI, 2007), como se ve reflejado en la teoría de Schumpeter.

A pesar de que las razones que llevan a la innovación en servicios e industria son similares, existen importantes diferencias entre ellas, siendo la principal la dificultad a la hora de distinguir entre productos y procesos en la primera (OECD y Eurostat, 2005; DTI, 2007). Esto provoca que la identificación de un proceso de innovación en servicios sea más trabajosa, e incluso no sea clasificada como innovadora por la empresa en la que tiene lugar (COTEC, 2007). Esta razón motiva la dificultad en la medición de innovación en servicios (OECD y Eurostat, 2005).

Dentro del sector servicios, los procesos de innovación son altamente heterogéneos, debido a las marcadas diferencias entre los distintos sectores, como distintos grados de cambio tecnológico o factores institucionales y organizacionales. Los servicios, y las innovaciones que pueden tener lugar en su seno, pueden clasificarse en cuatro categorías: servicios relacionados con bienes, relacionados con información, basados en el conocimiento y relacionados con las personas (OECD y Eurostat, 2005). No obstante, las innovaciones exitosas a menudo presentan características comunes, como la introducción en nuevos mercados (COTEC, 2005).

La innovación en servicios se divide a su vez en cuatro categorías (DTI, 2007) según el motor que lleve a ellas: originadas por los proveedores, por los clientes, innovaciones en servicios e innovaciones a través de servicios, pudiendo clasificarse en más de una categoría.

Si bien la mayor parte de las características de la innovación en el sector servicios son extensibles en cierto grado al resto de sectores, se considera que el estudio de la primera puede reportar mayores beneficios en un periodo de tiempo menor, ya que la innovación en servicios suele ser menos costosa y, por su actividad, se trata de empresas con mayor flexibilidad para alterar servicios o procesos. Es por ello que las conclusiones

de este trabajo serán de especial utilidad a estas empresas en su lucha por diferenciarse y alcanzar ventajas competitivas.

1.1.8. Innovación y ventaja competitiva

Como se ha enunciado en párrafos anteriores, la innovación es uno de los medios de los que se vale la empresa para alcanzar ventajas competitivas, es decir, obtener una rentabilidad superior a la de sus rivales (Grant, 2010) y generar valor.

La ventaja competitiva, consecuencia de la existencia de competencia imperfecta en los mercados, es un desequilibrio que surge como consecuencia de determinadas decisiones, y que tiende a desaparecer en el largo plazo, ya sea por imitación de los competidores o por desaparición de las condiciones externas que hacen que ese desequilibrio tenga lugar. Las decisiones tomadas por la empresa que provocan la aparición de ventajas competitivas pueden ser, por tanto, respuesta a un factor interno o una oportunidad externa.

La innovación es uno de los medios de los que puede servirse una entidad para alcanzar liderazgo en costes, por ejemplo a través de economías de escala, o en precios, por ejemplo ofreciendo productos de mayor precio a un nicho de mercado reducido en varios países. Gracias a la obtención de este liderazgo, la entidad consigue alcanzar ventajas competitivas, y con ello la obtención de mayores beneficios y rentabilidad. Este trabajo se centrará en intentar hallar los parámetros que favorecen el éxito en un proceso innovador, como *guía de ruta* a aquellas empresas que elijan emprender un proceso innovador para diferenciarse en el mercado.

1.1.9. Factores que llevan a la innovación

Las empresas llevan a cabo procesos de innovación con el fin último de mejorar su *performance* o resultados. El empresario puede decidir impulsar la innovación en casos en los que los costes de producción son altos, cuando la productividad es baja, o con el fin de aumentar la oferta o elevar el valor de mercado de la empresa (Rubera y Kirka, 2012). También pueden tener lugar factores externos, como pueden ser legales, u otros

motivos específicos a la empresa, como la escasez de trabajadores cualificados (OECD y Eurostat, 2005).

Entre los factores externos que pueden motivar la innovación empresarial, ocupa un importante lugar la protección que los gobiernos otorguen, generalmente en la forma de patentes, ya que los altos costes y el carácter incierto de la innovación actúan como una importante barrera de entrada (OECD y Eurostat, 2005). No obstante, un nivel elevado de protección puede dar lugar a incrementos en precios y a la interrupción al flujo de conocimientos que puede ser perjudicial a la economía en su conjunto.

Otros factores externos de gran importancia son la existencia de laboratorios y políticas estatales y universidades que favorezcan la innovación. Es por ello que en el proceso innovador cobra una especial importancia la cooperación entre las distintas partes (Aschoff y Schmidt, 2008). A pesar de ello la decisión de cooperar ha de ser tomada con precaución, ya que en casos los beneficios de la cooperación, tales como incremento de beneficios y acceso a conocimientos externos, deben superar el coste de la innovación.

1.1.10. Factores que influyen en el logro del éxito innovador

Para que un proceso innovador sea considerado de éxito, ha de traducirse en resultados para la empresa. Esta es la idea central sobre la que se desarrollarán los siguientes apartados: la identificación de perfiles de innovación de éxito a través de las distintas características que favorecen la innovación en el tejido empresarial.

En los siguientes sub-apartados se analizan los factores que a priori, y de acuerdo a la literatura, se consideran claves en el éxito de un proceso de innovación, con el fin de que actúen de aproximación a la segmentación por perfil de innovación. De esta forma se validará que estos factores actúan también de *potenciadores de éxito en innovación* para nuestro universo de estudio.

1.1.10.1. *Tamaño de la empresa*

De acuerdo a Zoltan y Audretsch (1987) el tamaño de la empresa es uno de los principales condicionantes del éxito en la innovación. Por un lado, las grandes empresas

tienden a poseer cierta ventaja competitiva, e incluso cierto poder monopolístico, en industrias que producen un bien muy diferenciado: intensivas en capital y con alta concentración. No obstante, las empresas pequeñas tienen ventaja competitiva en cuanto a éxito innovador en aquellos mercados que son intrínsecamente innovadores, con trabajadores altamente cualificados.

1.1.10.2. Edad

Otro de los factores que influye en el grado de innovación y en el éxito innovador de la empresa es su madurez, relacionado en cierta medida con el tamaño de la empresa. Según Wessel (2012), las empresas en sus primeros años están más centradas en la satisfacción de necesidades del mercado a través de la implementación de distintas formas de innovación, sin orientar su actividad exclusivamente a la búsqueda de rentabilidad para sus *stakeholders*, mientras que las empresas maduras tienen una mayor presión en la obtención de rentabilidad y eficiencia. De esta forma, mientras que las empresas maduras son más propensas a implementar innovaciones incrementales y a buscar eficiencia en sus procesos, las empresas jóvenes buscan soluciones innovadoras, conscientes de que si la encuentran verán por fin el crecimiento en sus ingresos.

1.1.10.3. Sector de actividad

Ha sido probado en numerosos estudios la vinculación del sector de actividad en el que está arraigada la empresa y su nivel de innovación, así como la probabilidad de éxito en los procesos innovadores que emprenda.

La razón del vínculo entre innovación y actividad se debe a la naturaleza del sector económico en el que la empresa realiza su actividad. Por su estructura de demanda y costes, las características del mercado concreto o el valor que en ese mercado se otorgue a la innovación (Flaig y Stadler, 1994).

Así mismo, juega un importante papel la concentración del mercado, que se tratará a lo largo de este trabajo, como medida del grado de competitividad del sector de actividad (Zoltan y Audretsch, 1987). Junto con la concentración, han de considerarse también las barreras de entrada que presente el sector a nuevos competidores, la posición

de la empresa en la industria y la heterogeneidad de ésta. Todas estas características favorecen la creación de un entorno que favorece la innovación y su éxito.

Junto con las características intrínsecas al sector de actividad, debe la valorarse la existencia de clústeres en los que se favorezca la innovación. En este caso entraría en juego no solo el sector de actividad, sino también la ubicación geográfica. Numerosos estudios certifican la asociación entre la existencia de clústeres e innovación (Baptista y Swann, 1998; Leifer y Colarelli, 2001; San Román et al, 2013).

1.1.10.4. Financiación Pública

Un alto porcentaje de la innovación se realiza amparada por el estado, como es el ejemplo de las Universidades o los Institutos de Investigación, o en el marco de determinadas políticas que la favorecen, tanto a nivel legislativo como fiscal. No obstante, como factor determinante del éxito innovador, tendremos únicamente en cuenta la Financiación directa del Estado.

Aunque a priori la existencia de financiación pública que favorezca la innovación resulta beneficioso, ha de analizarse detenidamente. En primer lugar, la eficiencia de la inversión y el grado en el que ésta se traduce en productos innovadores de los que se beneficie innovador y sociedad. También ha de considerarse si las ayudas públicas incrementan el nivel total de inversión en financiación, o si por el contrario meramente reemplazan a la inversión privada. Diferentes estudios empíricos han demostrado que la financiación pública únicamente tiene efectos *aditivos*, es decir, se traducen en incremento del gasto total en innovación, en el caso de las empresas pequeñas (Löf y Heshmati, 2004).

1.1.10.5. Gasto en innovación

Una característica importante pero no imprescindible para que tenga lugar un proceso innovador es que las empresas incurran en gastos en innovación. Pueden darse procesos innovadores en empresas que no dediquen parte de su presupuesto a innovación: ésta puede ser iniciativa de un departamento concreto como solución a un problema

determinado. Los gastos en innovación pueden ser realizados en el seno de la empresa (interna) o subcontratados a otras empresas (externa).

No obstante, y además de la importancia de este factor, también se ha probado la marcada influencia positiva que tiene haber realizado innovaciones en el pasado (Flaig y Stadler, 1994).

1.2. Base de Datos

Para analizar los procesos innovadores en las empresas españolas se ha utilizado la base de datos [Panel de Innovación Tecnológica \(PITEC\)](#), que se realiza anualmente el Instituto Nacional de Estadística (INE). Se trata de un instrumento estadístico que pretende facilitar el estudio de las innovaciones tecnológicas que tienen lugar en las empresas españolas, haciendo un especial énfasis en el estudio de las actividades de I+D emprendidas. La creación de esta base de datos surge como solución a la principal restricción que encuentran los investigadores a la hora de estudiar la innovación y el proceso innovador: la carencia de información fiable y homogénea acerca de las actividades innovadoras que llevan a cabo las empresas (Barajas, 2012, PITEC, 2012).

El PITEC es una base de datos de panel, con datos que abarcan un periodo comprendido entre 2003 y 2013 (última información disponible). Se ha diseñado según el marco armonizado de indicadores de innovación que marca el Manual de Oslo, cuyo contenido se ha descrito en apartados anteriores, y cuya primera edición se publicó en el año 1992.

La base de datos se compone de cuatro muestras:

- Muestra de empresas grandes (MEG), con un 73% de representatividad según evaluaciones del Directorio Central de Empresas (DIRCE) (PITEC, 2012).
- Muestra de empresas de cualquier tamaño que presentan gastos en innovación interna (MID).
- Muestra de empresas de menos de 200 empleados que incurren en gastos por compra de servicios de I+D, pero sin gastos de I+D interna (MIDE)

- Muestra de empresas de menos de 200 trabajadores sin gastos de innovación (MEP).

2. Objetivos

Se pretende realizar una segmentación de las empresas españolas según su perfil innovador, con el fin de identificar las características comunes a un proceso innovador de éxito. La segmentación se realizará en primer lugar a través de técnicas de aprendizaje no supervisado que aportarán la intuición necesaria para realizar la segmentación final, a través de un árbol de decisión.

Con el fin de analizar la influencia de la crisis económica española en los perfiles de innovación de las empresas, se realizarán dos segmentaciones para dos periodos de análisis: el periodo 2005-2007 y el periodo 2009-2011. Con ello se podrá realizar el estudio de los procesos innovadores antes y durante la crisis económica, así como sus posibles diferencias y analogías.

3. Metodología Empleada

3.1. Herramientas

Para realizar este trabajo se han empleado las siguientes herramientas:

- *Microsoft SQL Server*, como principal instrumento de gestión de la base de datos y el proceso ETL (Extracción, Transformación y Carga), a través de cual se ha realizado el estudio y auditoría de las variables y su comportamiento, la normalización y transformación de las variables existentes, así como la creación de nuevas variables sintéticas. Los *scripts* generados a través de esta herramienta pueden encontrarse en el anexo.

- *IBM SPSS Modeler* e *IBM SPSS Statistics*, como herramientas de modelación. Los *streams* generados a través de estas herramientas pueden encontrarse en los anexos.
- *Microsoft Excel*, como herramienta ocasional para la visualización de datos. Los análisis realizados que se consideran de mayor valor e interés se encuentran a lo largo del texto y sus correspondientes anexos.

3.2. Metodología

3.2.1 Proceso ETL

En primer lugar, tras recibir los archivos de datos correspondientes a cada año de realización de la encuesta en archivos de texto plano, se ha procedido a su carga en una base de datos de Microsoft SQL Server.

Las tablas brutas se normalizan e incluyen en un tablón de análisis que contiene la totalidad de la información, identificada de forma única por la clave de cliente y el año de realización de la encuesta. Para crear el tablón de análisis ha sido necesario completar las variables de los primeros años, de forma que para cada año se dispusiese de las 569 variables disponibles en 2013. Estas variables, que se introducen únicamente a efectos de homogeneización, se incluyen con valores en blanco o a NULL.

Una vez se dispone del tablón de análisis total, se procede a la creación de catálogos que permitan simplificar el análisis de las variables categóricas. Tras realizar diversos análisis y transformaciones a las variables de la base de datos que se especificarán en apartados posteriores, se crearon dos tablonas según la fecha de realización de la encuesta: tabla 2005-2007 (en adelante Tablón Pre) y tabla 2009-2011 (en adelante Tablón Post). Sobre estas tablas se realizarán los modelos estadísticos correspondientes en SPSS.

En un último paso, los resultados obtenidos de la segmentación en SPSS vuelven a introducirse en la base de datos de Microsoft SQL Server para continuar su análisis.

3.2.2. Segmentación

La segmentación es un procedimiento que consiste en asignar a los individuos del universo de análisis a un único grupo (clúster o segmento) de modo exhaustivo, es decir, todos los individuos pertenecen a un grupo, y mutuamente exclusivo, cada individuo pertenece a un único segmento. Así se trata de maximizar la homogeneidad de los individuos dentro de cada segmento y las diferencias entre segmentos.

Para producir estas agrupaciones las herramientas más habituales son los árboles de decisión y los algoritmos de *clustering*, cuya utilización dependerá de los objetivos que se persiga. En este trabajo se utilizarán dos: técnicas de aprendizaje supervisado y técnicas de aprendizaje no supervisado.

Una vez creados los distintos clúster, se procederá a caracterizarlos utilizando tanto variables características de innovación como variables características de la empresa, tales como el sector de actividad, el número de empleados, la madurez, la pertenencia a un grupo o los ingresos por ventas.

3.2.2.1. *Técnicas de aprendizaje supervisado*

Este tipo de segmentaciones pretende separar a los individuos en grupos homogéneos según una variable objetivo. Si este es el caso, se utilizan árboles de decisión como medio para explicar la variable objetivo, que es la que orienta el aprendizaje del modelo.

3.2.2.2. *Técnicas de aprendizaje no supervisado*

Los algoritmos de *clustering* trabajan sin supervisión, es decir, sin la guía de ninguna variable en particular. Todos los datos de los que se dispone entran al análisis, y el éxito consiste en agrupar a los individuos en segmentos que resulten significativos para los objetivos del negocio. Estos algoritmos pueden ser de fundamento estadístico o neuronal, como las redes de Kohonen o el algoritmo de *clustering two-step* o bifásico.

- Clustering bifásico

En este caso, se utilizará el *clustering bifásico*, un método de análisis de conglomerados de entrenamiento no supervisado en dos pasos. En el primer paso se hace una sola pasada a los datos, durante la cual se comprimen los datos de entrada sin procesar en un conjunto manejable de subclusters. El segundo paso utiliza un método de agrupación jerárquica para combinar progresivamente los subclusters sin necesidad de otro paso a través de los datos.

La agrupación jerárquica tiene la ventaja de no requerir el número de grupos que se seleccionarán antes de tiempo. Además, permite manejar variables discretas campo y es capaz de manejar grandes conjuntos de datos de forma eficiente. No obstante, el modelo resultante dependerá hasta cierto punto del orden de los datos de entrenamiento, por lo que la reordenación de los datos y la reconstrucción del modelo pueden conducir a un modelo de *clustering* diferente.

4. Desarrollo del Trabajo

4.1. Comprensión de los datos

Si bien los contenidos del cuestionario inicial se han mantenido constantes a lo largo del periodo de realización de la encuesta, se han ido añadiendo algunas variables, como puede observarse en la Tabla 1, junto con el número de empresas participantes y la tasa de participación efectiva:

Año	Número de Variables	Número de Registros	% Participación efectiva
2013	569	12.839	71%
2012	569	12.838	75%
2011	517	12.828	78%

Año	Número de Variables	Número de Registros	% Participación efectiva
2010	517	12.821	81%
2009	513	12.817	84%
2008	509	12.813	87%
2007	467	12.808	91%
2006	465	12.803	94%
2005	463	12.803	94%
2004	458	10.323	98%
2003	449	7.283	99%

Tabla 1 Composición de la Base de Datos PITEC

La base de datos, además de las variables que describen la innovación, tales como el gasto total en innovación o el porcentaje de bienes innovadores para el mercado que se producen, incluye otras variables descriptivas que permiten obtener una mejor comprensión de la empresa. Estas variables son, por ejemplo, el tamaño o la madurez de la compañía participante. Además de estas variables, se incluyen diversos indicadores que recogen las diversas incidencias que hayan podido tener lugar, como el cierre temporal o la regulación de empleo, con el fin de permitir la comparabilidad del conjunto de estudio.

En algunos casos, la información contenida en las variables ha sido anonimizada para evitar problemas de confidencialidad de los participantes. Para permitir la comparabilidad a lo largo del tiempo, se incluye un identificador único de empresa. En la Tabla 1 puede observarse el número de variables y registros, así como el porcentaje de participación efectiva en los años de realización de la encuesta, mientras que en la Tabla 2 se puede observar la continuidad en la participación de las empresas, según el año de realización de la encuesta.

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
2003	7.238	7.074	6.795	6.736	6.497	6.285	6.089	5.895	5.672	5.498	5.290
2004		10.093	9.636	9.614	9.262	8.935	8.619	8.295	7.969	7.671	7.347
2005			12.098	12.034	11.589	11.172	10.782	10.363	9.953	9.578	9.138
2006				12.034	11.589	11.172	10.782	10.363	9.953	9.578	9.138
2007					11.594	11.177	10.787	10.368	9.958	9.583	9.143
2008						11.182	10.792	10.372	9.962	9.587	9.147
2009							10.796	10.376	9.966	9.591	9.151
2010								10.380	9.970	9.595	9.155
2011									9.977	9.602	9.161
2012										9.612	9.171
2013											9.172

Tabla 2 Continuidad en la participación efectiva de las empresas

De las 569 variables de las que se dispone inicialmente, se ha reducido el número de variables que entrarán al tablón de análisis a través de diversas técnicas:

- Conversión de variables binarias en categóricas que sintetizan un conjunto de variables en una sola.
- Descarte de variables no relacionadas con el objeto de estudio.
- Descarte de variables que no están disponibles para todos los años que se analizan en este trabajo.
- Creación de ratios o variables sintéticas que agrupan variables y conservan la información de todas ellas.

Tanto el estudio del comportamiento de las variables, como su tratamiento (cambio de tipología de variable, análisis de valores *missing* o nulos, análisis de valores atípicos) se ha realizado de forma manual a través de Microsoft SQL Server, siempre buscando conservar sentido de negocio en los análisis. La sintaxis de estos análisis, así como la sintaxis de normalización y de creación de las distintas tablas puede encontrarse en el [Anexo VII](#).

4.2. Selección de fechas de referencia

Como se ha descrito en apartados anteriores, se han creado dos tabloneros – Tablón Pre y Tablón Post – cuyos periodos considerados son 2005-2007 y 2009-2011

respectivamente. No obstante, se ha elegido de fecha de referencia el último año de cada periodo, 2007 y 2011, con el fin de simplificar el análisis, y se calcularán las variaciones contra el promedio de los dos primeros años del periodo a considerar.

4.3. Creación de variables sintéticas preliminares

Con el fin de facilitar el análisis de las 569 variables, se han creado diversas variables sintéticas.

4.3.1. Tipo de empresa

Las empresas se han clasificado en cuatro tipos o tamaños, según su número de empleados y de acuerdo a la [Legislación de la Unión Europea](#):

- *Microempresa*, con 10 o menos empleados
- *Pequeña empresa*, con un máximo de 50 empleados
- *Mediana empresa*, con un máximo de 250 empleados
- *Gran empresa*, con más de 250 trabajadores

Además sea comprobado que los resultados obtenidos según este criterio son aproximadamente iguales a los obtenidos a través de la distribución de las empresas en cuartiles según su número de empleados.

4.3.2. CNAE

Ante el cambio de las correspondencias de la clasificación de códigos de CNAE correspondiente a CNAE-93 (Tablón Pre) y CNAE-2009 (Tablón Post), se ha creado una equivalencia de acuerdo al Instituto Nacional de Estadística para mantener la comparabilidad entre ambos periodos.

Además, para simplificar el análisis, se ha realizado una agrupación de CNAEs según su sentido de negocio en las siguientes categorías:

- *Comercio*
- *Construcción*
- *Fabricación*

- *Ingeniería*
- *Investigación*
- *Manufacturas*
- *Materias Primas*
- *Tecnología*
- *Transporte y Motor*
- *Servicios*
- *Servicios primarios*

El detalle del análisis realizado para diseñar esta agrupación podrá encontrarse en el [Anexo I](#) y el [Anexo II](#).

4.3.3. Concentración de la industria

Con el fin de medir el nivel de competitividad del mercado se ha introducido la variable Concentración de la Industria a través del cálculo del Índice de Herfindahl e Hirschman (IHH). Este índice se calcula mediante la suma de los cuadrados de la cuota de mercado de cada empresa de la industria, donde valores elevados indican un mercado poco competitivo.

En nuestro universo de datos, se ha calculado un valor para cada CNAE, obteniéndose mayores más elevados del índice en industrias como el Tabaco o la industria de Refino de Petróleo. Se comprueba así que de acuerdo con la intuición, se trata de mercados muy concentrados y poco competitivos, muy cercanos al monopolio/oligopolio.

En pasos posteriores se ha clasificado esta variable continua en 5 categorías para facilitar el análisis. El detalle, tanto del IHH como de su categorización para CNAE-93 y CNAE-2009, puede encontrarse en el [Anexo III](#) y el [Anexo IV](#).

4.3.4. Categorización de madurez

La madurez de la empresa, que se ha obtenido a través del año de fundación de la empresa (disponible a partir del año 2009) y del año de referencia, se ha dividido en seis categorías:

- Empresas con menos de 5 años
- Empresas de entre 5 y 12 años

- Empresas de entre 12 y 20 años
- Empresas de entre 20 y 30 años
- Empresas de más de 30 años
- Empresas sin madurez informada

4.3.4. Ratio de innovación y volumen de negocio

El ratio de innovación y volumen de negocio se calcula mediante la división entre el gasto total en innovación y el volumen de negocio en el año de referencia.

A través de este ratio se pretende conocer qué porcentaje de sus ingresos por ventas dedican las empresas a la innovación, como indicador de cómo de rentable es la innovación a la empresa, y así de esta forma suplir la carencia de variables exclusivas de rentabilidad en el conjunto de datos.

4.3.5. Variaciones o Deltas

Se han calculado variaciones contra el promedio de los dos primeros años de cada intervalo considerado para las siguientes variables:

- Gasto total en innovación
- Volumen de negocio
- Ratio innovación y volumen de negocio
- Número de empleados
- Porcentaje de productos nuevos para el mercado
- Porcentaje de productos nuevos para la empresa

4.3.6. Otras variables

Se han creado otras variables como el índice de concentración, que categoriza la variable concentración de la industria en cinco intervalos, según su grado de competitividad.

También se han creado otras variables, que agrupan variables binarias del conjunto de datos original en una única variable categórica.

4.4. Definición del Público Objetivo

Con el fin de mantener la comparabilidad y facilitar la comprensión y análisis de los resultados obtenidos, se han definido las siguientes condiciones para las empresas que entran a formar parte del público objetivo:

- Empresas participantes en los periodos de análisis considerados.
- Empresas que no presentan incidencias del tipo:
 - Cierre temporal
 - Cierre definitivo
 - Contenida en otra
 - Erróneamente incluida
 - Absorción
 - Fusión
 - Duplicidad
 - Escisión total
 - Ilocalizable
- Empresas con volumen de negocio mayor que cero en la fecha de referencia.

Con estas condiciones se obtienen 11.585 empresas en el primer periodo y 9.949 empresas en el segundo periodo. Esto supone respectivamente un 90,45% y un 77,56% de la totalidad de individuos que entran a la encuesta, siendo un primer indicador de los efectos de la crisis económica en las empresas españolas.

El detalle y la sintaxis de creación de los dos tablonos de análisis puede encontrarse en los *scripts* 20151031_CREACION TABLON ANALISIS PRE_V2 y 20151101_CREACION TABLON ANALISIS POST_V1 del [Anexo VII](#), junto con el resto de scripts.

4.5. Creación de variable objetivo

Ante la carencia de variables relacionadas con la rentabilidad empresarial, o variable de las cuales ésta pudiera derivarse, se ha diseñado una variable objetivo ad-hoc

que describe el éxito en el proceso innovador a través de las variables Volumen de Negocio, Número de Empleados y diversas variables relacionadas con la innovación.

Se considera entonces que una empresa lleva a cabo un proceso innovador de éxito cuando se cumplen las siguientes condiciones:

- El gasto total en innovación en los dos primeros años de cada periodo es mayor que cero.
- Ha aumentado el número de empleados o el volumen de negocio en el año de referencia con respecto al promedio de los dos años anteriores.
- Se cumple por lo menos una de las siguientes condiciones:
 - El gasto total en innovación en el año de referencia es mayor que cero,
 - Hay una innovación en curso,
 - Se han llevado a cabo innovaciones de bienes, procesos, productos o servicios en los dos años anteriores.

Con estas condiciones se obtienen:

- 380 éxitos en el primer periodo, sobre un total de 11.585 empresas.
- 98 éxitos en el segundo periodo, sobre un total de 9.949 empresas.
- 15 éxitos comunes a ambos periodos.

4.6. Segmentación de empresas 2005-2007

Tras la obtención del Tablón Pre, se procede a su conexión con IBM SPSS Modeler para su tratamiento previo a la modelación. Una vez realizados los análisis y clasificado las variables nuevamente según su tipo, entran al análisis las siguientes variables:

Variable	Descripción
IDENT	Identificador de empresa
CNAE_07	Código CNAE 93
CATEGORIA_CNAE	Agrupación de CNAEs en 10 categorías
CONCENTRACION	Índice de concentración de la industria (1- oligopolio)
MUESTRA	Muestra a la que pertenece la empresa (

Variable	Descripción
IND_CONCENTRACION	Clasificación de concentración de la industria (1- oligopolio)
IND_BIOTECNOLOGIA	Indicador de realización de biotecnología
IND_FIN_PUBLICA	Indicador de financiación pública
DESC_CLASE	Clase de empresa: Asociación/institución, Pública, Privada Nacional, Privada Multinacional
GRUPO	Indicador de pertenencia a grupo
NUMEMP	Número de empleados
TIPO_EMPRESA	Tamaño de la empresa según su número de empleados: Micro, Pequeña, Mediana, Grande
DELTA_NUMEMP_0_1_2T	Variación del número de empleados con respecto al promedio de los dos últimos años (crecimiento si > 1)
MADUREZ	Años de la empresa
CLASIF_MADUREZ	Clasificación de años de la empresa
VOLNEG	Volumen de negocio
DELTA_VOLNEG_0_1_2T	Variación del volumen de negocio con respecto al promedio de los dos últimos años (crecimiento si > 1)
G_TOTAL_INN	Gastos Totales en innovación
RATIO_GINN_VOLNEG	Ratio Gastos en innovación / Volumen de negocio
DELTA_RATIO_GINN_VOLNEG_0_1_2T	Variación del ratio Gastos en innovación / Volumen de negocio con respecto al promedio de los dos últimos años
DELTA_G_TOTAL_INN_0_1_2T	Variación del gasto en innovación con respecto al promedio de los dos últimos años (crecimiento si > 1)
INN_0T	Innovación en curso
INN_PROD_2T	Innovación de producto hace dos años
INN_BIEN_2T	Innovación de bienes hace dos años
INN_SERV_2T	Innovación de servicios hace dos años
INN_PROCESO_2T	Innovación de proceso hace dos años
TIPO_INNOVACION	Innovación con respecto a: Empresa, Mercado, Empresa y Mercado, Sin Información
TIPO_INNOVACION_PROC_2Y	Tipo de innovación de proceso hace dos años
IND_GASTOS_EXT_INN	Indicador de gastos externos en innovación
IND_GASTOS_INT_INN	Indicador de gastos internos en innovación
DELTA_PROM_NUEVOEMPR_0_1_2T	Variación del % de productos innovadores para la empresa con respecto al promedio de los dos últimos años (crecimiento si >1)
DELTA_PROM_NUEVOMERC_0_1_2T	Variación del % de productos innovadores para el mercado con respecto al promedio de los dos últimos años (crecimiento si >1)
IND_PATENTES	Indicador de solicitud de patentes
IND_GASTOS_MAQUI	Indicador de gastos en maquinaria
IND_GASTOS_TECNO	Indicador de gastos en tecnología
IND_GASTOS_PREP	Indicador de gastos en preparación del producto
IND_GASTOS_FORM	Indicador de gastos en formación
IND_EXITO	Indicador de éxito

Tabla 3 Variables que entran al análisis

El conjunto de datos tratado en SPSS se exporta nuevamente a IBM SPSS Statistics, donde se procederá a su modelación.

4.6.1. Análisis no jerárquico

Con el fin de determinar el número óptimo de conglomerados o clúster que deben obtenerse y dado el volumen de datos a analizar, se aplicará el método de clustering bifásico o two-step. Ambos algoritmos se han descrito en apartados anteriores.

A través de sucesivas iteraciones, y partiendo de diez conglomerados, se unen grupos similares con centroides cercanos hasta obtenerse que el número óptimo de clústeres es de 4. El resumen del modelo bietápico de tres conglomerados puede verse en la imagen inferior:

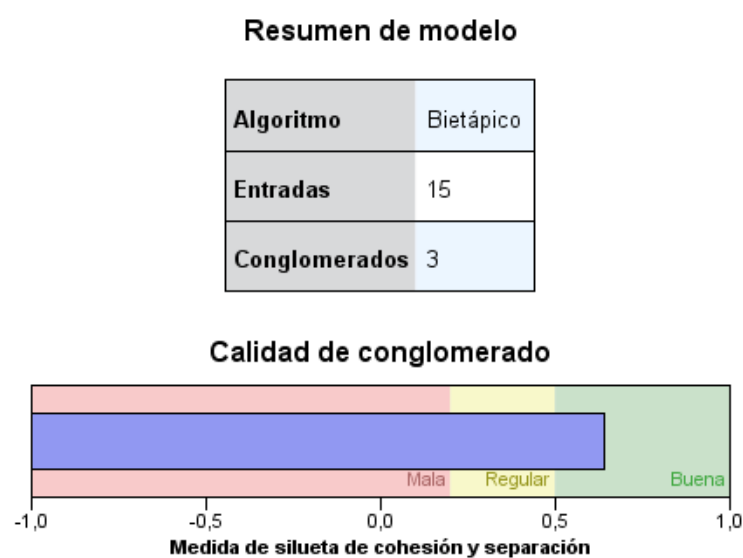


Gráfico 1 Salida SPSS Bietápico

Conglomerados		
1	1.304,000	11%
2	5.350,000	46%
3	4.931,000	43%

Tabla 2 Distribución de Observaciones por Conglomerado

Casos Válidos	11.585,000	100%
Casos Perdidos	0,000	0%

Tabla 3 Cantidad de registros analizados

Se observa que el número de registros de cada conglomerado es representativo, por lo que se mantienen los tres conglomerados. No se ha considerado necesario aplicar otro método de análisis no jerárquico, por obtenerse resultados coherentes, representativos y con sentido de negocio. A continuación se describen las principales características de los clústeres obtenidos:

- Clúster 1: Empresas de sectores económicos de baja concentración económica que presentan niveles moderados de innovación.
- Clúster 2: Empresas de sectores económicos de alta concentración, que no presentan actividades innovadoras en el periodo considerado.
- Clúster 3: Empresas de sectores económicos de alta concentración, que presentan niveles altos de innovación.

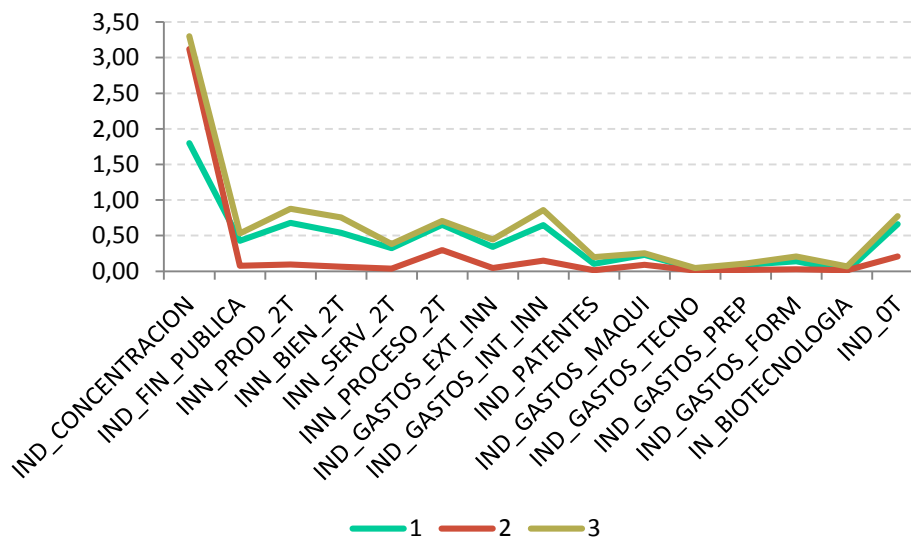


Gráfico 2 Centros de los conglomerados

4.6.2. Árbol de clasificación

Una vez obtenidos los clústeres de empresas del análisis no jerárquico, y tomando esta información como input, se procede a aplicar un árbol de decisión que permita predecir la variable objetivo seleccionada: el indicador de éxito en el proceso de innovación según el incremento de ventas.

Se ha utilizado el árbol de decisión CHAID (Detector automático de Chi-cuadrado de interacción), que permite realizar varias divisiones en cada rama. A pesar de que IBM SPSS Modeler permite la generación de árboles de decisión de manera automática, se ha

decidido hacerlo de manera manual, de forma que se tenga control sobre las variables que segmenten al público objetivo. Con ello podremos identificar qué segmentos presentan un mayor índice de éxitos en innovación, lo que permitirá dar a conocer *best practices* o estrategias de éxito de las que podrá beneficiarse la sociedad en su conjunto.

El árbol de decisión obtenido para el que se producen las mejores ganancias y que además presenta estabilidad en el conjunto de entrenamiento y test con distintas semillas aleatorias, divide a las empresas participantes en 15 segmentos diferentes, con un mínimo de 100 individuos en cada segmento. La representación gráfica de este árbol puede encontrarse en el [Anexo V](#). Las variables utilizadas en la segmentación que sea encontrado que más discriminan la variable objetivo o target han sido:

- Tipo de Empresa, siendo las más innovadoras las Microempresas
- Clasificación de su sector de actividad, siendo las más innovadoras las empresas de Ingeniería e Investigación
- Madurez de la Empresa, siendo más innovadoras en sus primeros años
- Indicador de Financiación Pública
- Indicador de Gastos Internos de Innovación

De esto modo, las mayores mejoras sobre la tasa de éxito inicial o prior (3,576%) se obtienen para empresas de Investigación o Ingeniería con menos de 12 años de madurez y que reciben financiación pública, o, en general, para microempresas con menos de 12 años de madurez que presentan gastos en innovación. Los segmentos obtenidos se caracterizan de la forma:

- *Segmento 1*, empresas jóvenes de Investigación o Ingeniería que reciben financiación pública. Es el segmento con mayor ratio de éxito, y el que presenta el mayor ratio de Gastos de Innovación sobre Volumen de Negocio. En su mayor parte se trata de microempresas y pequeñas empresas nacionales privadas con innovaciones en curso. Estas empresas incurren en gastos de innovación interna, en gastos de maquinaria y en

gastos de formación, y presentan el mayor porcentaje de solicitud de patentes. Este segmento tiene el ratio más elevado de éxitos en el segundo periodo.

- *Segmento 2*, empresas jóvenes de Investigación o Ingeniería que no reciben ayudas públicas, con un mayor porcentaje de éxitos que el prior de la cartera, aunque dedican una mínima proporción de su volumen de negocio a los gastos en innovación. Se trata en su mayor parte de microempresas o pequeñas empresas con un volumen de negocio relativamente pequeño.
- *Segmento 3*, empresas consolidadas de Investigación o Ingeniería que reciben ayudas públicas. Se trata en su mayor parte de empresas pequeñas y medianas, que han realizado innovaciones de productos y procesos en los dos años anteriores al año de referencia, y que presentan un ratio de Gastos de Innovación sobre Volumen de Negocio relativamente elevado. Incurren en gastos de formación.
- *Segmento 4*, empresas consolidadas de Investigación o Ingeniería que no reciben ayudas públicas. Presenta un ratio de Gastos de Innovación sobre Volumen de Negocio muy elevado. Tienen innovaciones en curso pero la mayor parte no incurre en gastos por innovación de ningún tipo, por lo que su tasa de éxitos es la misma que la cartera.
- *Segmento 5*, empresas jóvenes de Construcción o Tecnología que reciben ayudas públicas. Es uno de los tres segmentos con mayor ratio de éxito. Han realizado innovaciones de producto en el pasado, y tienen innovaciones en curso. Presentan gastos en formación.
- *Segmento 6*, empresas jóvenes de Construcción o Tecnología que no reciben ayudas públicas, y con un perfil mediocre de innovación. No han realizado innovaciones en el pasado e incurren en pocos gastos en innovación.
- *Segmento 7* empresas consolidadas de Construcción o Tecnología, la mayor parte de las cuales pertenece a un grupo. Han realizado

innovaciones de producto en años anteriores, pero invierten un volumen muy reducido de su volumen de negocio en gastos en innovación, lo que se traduce en un ratio de éxito reducido.

- *Segmento 8*, microempresas jóvenes del resto de sectores económicos que reciben ayudas públicas. Es uno de los tres segmentos con mayor ratio de éxito. Presenta un ratio de Gastos de Innovación sobre Volumen de Negocio muy elevado. Incurren en gastos por innovación interna y adquisición de maquinaria.
- *Segmento 9*, microempresas jóvenes del resto de sectores económicos que no reciben ayudas públicas. No han realizado innovaciones en años anteriores, y ni tienen innovaciones en curso.
- *Segmento 10*, pequeñas empresas jóvenes del resto de sectores económicos que presentan gastos en innovación interna.
- *Segmento 11*, pequeñas empresas jóvenes del resto de sectores económicos que no presentan gastos en innovación interna ni tienen innovaciones en curso.
- *Segmento 12*, medianas o grandes empresas jóvenes del resto de sectores económicos que reciben ayudas públicas. Se trata de empresas privadas nacionales o internacionales, la mayor parte de las cuales pertenece a un grupo, y que dedican un porcentaje reducido de sus ingresos a la innovación. Tienen procesos de innovación en curso.
- *Segmento 13*, medianas o grandes empresas jóvenes del resto de sectores económicos que no reciben ayudas públicas. Es uno de los tres segmentos con menor ratio de éxitos. Se trata de empresas privadas nacionales o internacionales que apenas realizan inversión en innovación.
- *Segmento 14*, empresas pequeñas o medianas consolidadas del resto de sectores económicos que no reciben ayudas públicas. Es uno de los tres segmentos con menor ratio de éxitos. Se trata de empresas privadas nacionales o internacionales que dedica un porcentaje muy reducido de sus

ingresos a la innovación. A pesar de ello, tienen innovaciones en curso, y han realizado procesos de innovación en el pasado.

- *Segmento 15*, empresas consolidadas del resto de sectores económicos que no reciben ayudas públicas. Es uno de los tres segmentos con menor ratio de éxitos. Se trata de empresas privadas nacionales o internacionales

4.6.3. Resultados

Durante el proceso de elaboración de los árboles de decisión ha sido puesta de manifiesto la escasa influencia que tienen sobre el target variables como la pertenencia a un grupo o el indicador de realización de innovación externa. Por otra parte, las empresas beneficiarias de Financiación Pública multiplican por dos su probabilidad de realizar procesos de innovación exitosos. No obstante, este indicador puede llevar a error, ya que puede darse el caso de que las ayudas públicas tengan como requisito para su recepción la realización de actividades o investigación en innovación.

La segmentación obtenida a través del árbol de clasificación concreta las intuiciones obtenidas de los clústeres de los algoritmos no jerárquicos, como son que las empresas que reciben financiación pública efectivamente tienen procesos de innovación en curso, y con un mayor porcentaje de éxito, o que las empresas con gastos internos en innovación también presentan innovaciones en curso en el año de referencia.

Con todo, ambas segmentaciones son complementarias, ya que la segmentación procedente del árbol de clasificación incluye variables más *sociodemográficas* y con ella se pretende conocer mejor cuáles son las empresas que realizan procesos innovadores de éxito, mientras que con los algoritmos no jerárquicos se ha pretendido generar distintos perfiles de innovación, sin entrar a su éxito.

4.7. Segmentación de empresas 2009-2011

A continuación se repite el análisis realizado en el apartado anterior para el Tablón Post, que considera los datos de las empresas participantes en el intervalo 2009-2011. Las variables de entrada son las mismas que en el caso del análisis del periodo 2005-2007.

4.7.1. Análisis no jerárquico

Partiendo otra vez de diez conglomerados, se obtiene que el número óptimo de clústeres es de 2. El resumen del modelo bietápico de tres conglomerados puede verse en la imagen inferior:

Resumen de modelo

Algoritmo	Bietápico
Entradas	14
Conglomerados	2

Calidad de conglomerado

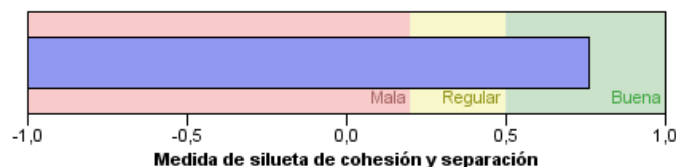


Gráfico 3 Salida SPSS Bietápico

Conglomerados

1	5.879,000	59%
2	4.073,000	41%

Tabla 6 Distribución de Observaciones por Conglomerado

Casos Válidos	9.949,000	100%
Casos Perdidos	0,000	0%

Tabla 7 Cantidad de registros analizados

Se observa que el número de registros de cada conglomerado es representativo, por lo que se mantiene el número de conglomerados. No se ha considerado necesario aplicar otro método de análisis no jerárquico, por obtenerse resultados coherentes, representativos y con sentido de negocio. A continuación se describen las principales características de los clústeres obtenidos:

- Clúster 1: Empresas que han llevado a cabo procesos de innovación en el intervalo considerado, más propensas a recibir financiación pública.
- Clúster 2: Empresas que no llevan a cabo procesos de innovación.

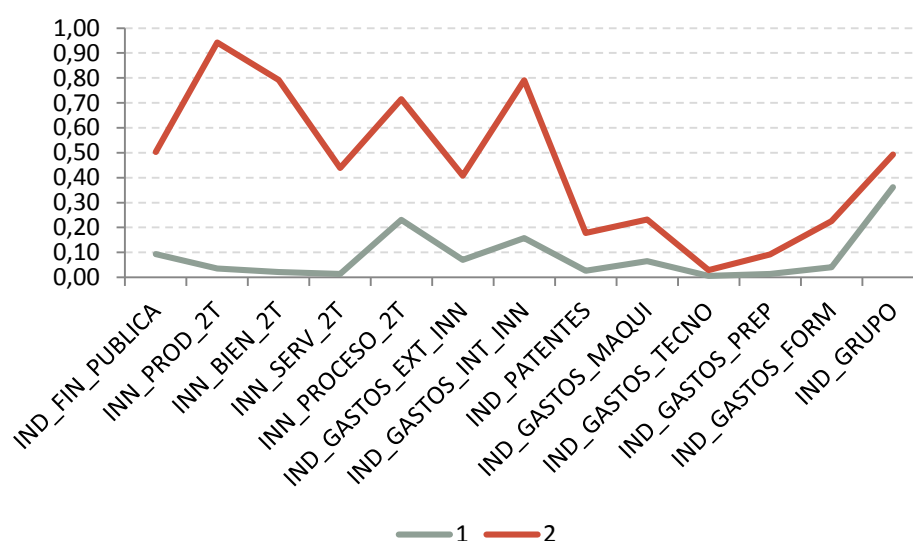


Gráfico 4 Centros de los conglomerados

4.7.2. Árbol de clasificación

La metodología utilizada para general el árbol de decisión en el periodo 2009-2011 es análoga a la utilizada en el periodo 2005-2007, y que se explica en el apartado 4.6.2 de este documento.

El árbol de decisión obtenido para el que se obtienen las mejores ganancias y que además presenta estabilidad en el conjunto de entrenamiento y test con distintas semillas aleatorias, divide al público objetivo en 9 segmentos diferentes, con un mínimo de 100 individuos en cada segmento. La representación gráfica de este árbol puede encontrarse en el [Anexo VI](#). Las variables utilizadas en la segmentación que sea encontrado que más discriminan la variable objetivo o target han sido:

- Clasificación de su sector de actividad, siendo las más innovadoras las empresas de Ingeniería e Investigación
- Madurez de la Empresa, siendo más innovadoras en sus primeros años
- Indicador de Financiación Pública
- Indicador de Gastos Internos de Innovación

En este segundo periodo de análisis las mayores mejoras sobre el prior inicial se obtienen, de forma significativa, para empresas clasificadas en los sectores de Investigación e Ingeniería, así como aquellas empresas que reciben financiación pública.

4.7.3. Resultados

En la segunda segmentación, como también sucede en el primer periodo, sigue siendo de relevancia la reducida importancia que tienen sobre el target variables que a priori pueden parecer importantes como la pertenencia a un grupo de empresas o el indicador de realización de innovación externa. Además, en este segundo periodo, variables que en el primero sí cobraban importancia, como el tamaño de la empresa, dejan de ser relevantes.

La segmentación obtenida a través del árbol de clasificación es reflejo de las intuiciones obtenidas de los dos clústeres de los algoritmos no jerárquicos. Puede apreciarse con facilidad como se agudiza la condición de ser beneficiarias de ayudas públicas para llevar a cabo un proceso innovador de éxito, o cómo las empresas que incurren en gastos de financiación internos son más innovadoras. Los segmentos obtenidos se caracterizan de la forma:

- *Segmento 1*, empresas jóvenes de Investigación, Tecnología o Ingeniería que reciben financiación pública, que presentan mayores mejoras en el ratio de éxitos con respecto al prior de la cartera. En su mayor parte son micro o pequeñas empresas que han llevado a cabo procesos de innovación de producto en el pasado, y con el mayor ratio de inversión en innovación sobre el volumen de negocio. Todas incurren en gastos de innovación interna e incluso de innovación externa, así como en gastos de formación.

- *Segmento 2*, empresas jóvenes de Investigación, Tecnología o Ingeniería que no reciben ayudas públicas. Invierten un porcentaje muy reducido de sus ingresos en innovación, hay llevado a cabo procesos de innovación muy modestos en los dos años anteriores y no incurren en gastos de maquinaria, tecnología o formación, por lo que el ratio de éxitos es el mismo que el prior de la cartera.
- *Segmento 3*, empresas consolidadas de Investigación, Tecnología o Ingeniería que no incurren en gastos de innovación interna. Invierten un porcentaje muy reducido de sus ingresos en innovación, y no incurren en ningún otro tipo de gastos. Tampoco realizan procesos de innovación en el pasado. Es uno de los tres segmentos con menor ratio de éxitos.
- *Segmento 4*, empresas consolidadas de Investigación, Tecnología o Ingeniería que incurren en gastos de innovación interna y reciben ayudas del Estado. En su mayor parte, también incurren en gastos de innovación externa y han realizado procesos de innovación de bienes y procesos en los dos años anteriores. Es uno de los tres segmentos que más porcentaje de su volumen de negocio invierte en innovación, a pesar de ello este porcentaje es reducido, por lo que el ratio de éxito es solo ligeramente superior al prior de la cartera.
- *Segmento 5*, empresas consolidadas de Investigación, Tecnología o Ingeniería que incurren en gastos de innovación interna y que no reciben ayudas del Estado. Incurren en gastos de formación y han realizado procesos de innovación en el pasado. No obstante, invierte un volumen muy reducido de sus ingresos en innovación, por lo que su ratio de éxito es solo ligeramente superior al prior de la cartera.
- *Segmento 6*, empresas medianas o grandes jóvenes del resto de actividades económicas que incurren en gastos en innovación interna, e incluso de innovación externa. Han realizado procesos de innovación de bienes y servicios en los dos últimos años, e incurren en gastos de maquinaria y formación. El ratio de éxitos está por encima del prior de la cartera.

- *Segmento 7* empresas jóvenes del resto de actividades económicas que no incurren en gastos en innovación interna. No han realizado procesos de innovación en el pasado y no tienen apenas gastos en innovación. Es por ello que es uno de los tres segmentos con menor ratio de éxitos.
- *Segmento 8*, empresas maduras del resto de actividades económicas que incurren en gastos en innovación interna. Han realizado innovaciones de bienes o procesos en los dos años anteriores, e incurren en gastos en maquinaria y formación. No obstante, invierten un porcentaje muy reducido de su
- *Segmento 9*, empresas maduras del resto de actividades económicas que no incurren en gastos en innovación interna. No han realizado procesos de innovación en el pasado y no tienen apenas gastos en innovación, ni gastos de maquinaria, tecnología o formación. Es uno de los tres segmentos con menor ratio de éxitos.

4.8. Comparativa

En la siguiente tabla se muestra la evolución de los segmentos definidos en el periodo 2005-2007, si esa segmentación volviese a aplicarse en el periodo 2009-2011. Puede observarse que la mayor parte de las empresas tiene un comportamiento estable.

Empresas	2009-2011																
2005-2007	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	NULL	Total general
1	47,27%	13,18%	16,08%	4,50%	2,25%	0,32%	0,64%	0,64%	0,64%	0,32%	0,32%	0,96%	0,00%	0,32%	0,32%	12,22%	100,00%
2	11,67%	32,78%	6,11%	26,67%	0,56%	4,44%	1,11%	0,00%	1,67%	0,56%	0,00%	0,00%	0,56%	0,00%	0,00%	13,89%	100,00%
3	0,00%	0,00%	54,44%	17,37%	0,00%	0,00%	3,09%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	1,54%	2,32%	21,24%	100,00%
4	0,00%	0,46%	10,14%	56,68%	0,00%	0,00%	4,15%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	3,69%	24,88%	100,00%
5	0,81%	0,54%	0,81%	0,81%	27,64%	22,22%	32,52%	0,54%	0,81%	0,27%	0,54%	0,00%	0,00%	1,36%	0,54%	10,57%	100,00%
6	0,00%	0,78%	0,19%	0,19%	6,63%	31,38%	35,67%	0,19%	1,36%	0,39%	0,78%	0,19%	0,58%	0,19%	5,85%	15,59%	100,00%
7	0,00%	0,00%	0,19%	0,71%	0,06%	0,06%	69,43%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	1,93%	6,68%	20,94%	100,00%
8	7,10%	5,81%	1,29%	1,94%	2,58%	3,23%	2,58%	10,32%	26,45%	4,52%	3,87%	0,00%	0,00%	7,10%	14,84%	8,39%	100,00%
9	0,79%	9,06%	0,00%	1,97%	0,39%	3,15%	3,94%	3,15%	33,86%	1,97%	3,94%	0,00%	0,39%	3,15%	20,08%	14,17%	100,00%
10	2,60%	1,30%	1,04%	2,60%	2,08%	3,38%	4,16%	1,82%	4,42%	15,06%	11,95%	2,08%	1,30%	11,43%	25,71%	9,09%	100,00%
11	1,06%	2,12%	0,00%	0,53%	0,53%	0,53%	1,33%	0,53%	6,10%	2,92%	18,30%	0,27%	1,86%	3,98%	45,62%	14,32%	100,00%
12	1,62%	1,08%	0,54%	0,54%	0,00%	1,08%	1,08%	0,00%	0,00%	1,62%	3,24%	22,70%	14,05%	25,95%	17,84%	8,65%	100,00%
13	0,16%	1,25%	0,00%	1,56%	0,16%	2,18%	2,03%	0,00%	0,47%	0,78%	1,56%	4,84%	30,89%	3,74%	42,59%	7,80%	100,00%
14	0,00%	0,00%	1,08%	1,01%	0,00%	0,00%	4,50%	0,00%	0,00%	0,00%	0,06%	0,00%	0,06%	40,03%	40,66%	12,60%	100,00%
15	0,00%	0,00%	0,22%	1,91%	0,00%	0,00%	3,74%	0,00%	0,02%	0,00%	0,00%	0,00%	0,04%	9,41%	68,91%	15,75%	100,00%
Total general	1,74%	1,40%	2,29%	3,28%	1,39%	2,56%	14,66%	0,33%	1,61%	0,81%	1,34%	0,74%	2,11%	10,84%	39,84%	15,06%	100,00%

Tabla 8 Movimiento de empresas según segmentación 2005-2007

Empresas	2009-2011										
2005-2007	1	2	3	4	5	6	7	8	9	NULL	Total general
1	39,14%	18,50%	5,66%	16,21%	5,50%	1,22%	1,22%	1,07%	0,31%	11,16%	100,00%
2	9,27%	34,44%	21,85%	4,30%	7,62%	0,66%	2,81%	1,49%	2,98%	14,57%	100,00%
3	0,00%	0,00%	51,74%	6,79%	5,40%	0,00%	0,00%	1,22%	9,93%	24,91%	100,00%
4	0,00%	0,00%	14,61%	49,51%	10,71%	0,16%	0,00%	3,90%	1,46%	19,64%	100,00%
5	0,00%	0,20%	22,56%	13,82%	33,13%	0,00%	0,00%	4,88%	4,67%	20,73%	100,00%
6	3,76%	4,73%	2,26%	1,72%	1,51%	24,30%	16,34%	23,12%	13,76%	8,49%	100,00%
7	0,93%	4,23%	3,13%	0,08%	0,17%	4,31%	31,30%	3,30%	40,19%	12,35%	100,00%
8	0,00%	0,00%	2,11%	2,21%	1,87%	0,00%	0,07%	56,81%	23,94%	13,01%	100,00%
9	0,00%	0,00%	4,26%	0,20%	0,28%	0,00%	0,11%	7,92%	70,23%	17,00%	100,00%
Total general	3,09%	3,66%	8,11%	5,46%	3,65%	2,50%	4,77%	19,70%	33,98%	15,06%	100,00%

Tabla 9 Movimiento de empresas según segmentación 2009-2011

Entre los 15 éxitos comunes a ambos periodos, destacan las siguientes características:

- Empresas privadas nacionales, micro o pequeñas
- Son beneficiarias de ayudas estatales
- Presentan un ratio elevado de inversión en innovación con respecto a su volumen de negocio
- Tienen innovaciones en curso e incurren en gastos de innovación interna, y en su mayor parte, de innovación externa
- La mayoría pertenecen a los sectores económicos de investigación, ingeniería y tecnología

5. Conclusiones

Con la realización de este trabajo se han podido confirmar, para nuestro universo de estudio, algunas de las intuiciones expuestas en el inicio. A continuación se enunciarán las principales conclusiones obtenidas, así como las principales pautas para facilitar o favorecer el éxito innovador de aquellas empresas que decidan emprender un proceso de innovación:

- Se ha observado el aumento generalizado de incidencias (fusión, cierre...) en el segundo periodo de estudio. Éstas se duplican en el periodo 2009-2011 con respecto al periodo 2005-2007

	Total Empresas	Total Incidencias	% Incidencias
PRE	12.808	513	4,01%
POST	12.828	1.040	8,11%

Tabla 10 Porcentaje de Incidencias

Así mismo, se produce también un descenso en el ratio de éxito innovador (prior) en el segundo periodo con respecto al primero:

	PRE	POST
Total Empresas	11585	9949
Total Éxitos	380,0	98,0
Ratio de éxito	3%	1%

Tabla 11 Ratio de éxito

NOTA: Se debe recordar que la diferencia entre el Total Empresas disponibles y el Total Incidencias no es el Público Objetivo, si no que se descartan las empresas que no informan o participan en la encuesta. Si se tiene en cuenta estas empresas, el porcentaje se dispara.

- Así mismo, se observa también un descenso del ratio de inversión en el segundo periodo con respecto al primero, a pesar de que el gasto medio en innovación permanece constante.

	PRE	POST
Total Empresas	11.585	9.949
Ratio innovación	77,1	71,6
Media gasto en innovación	982.229€	1.056.012€

Tabla 12 Innovación por periodo

- De los resultados de las segmentaciones se confirma la lógica: incurrir en gastos de innovación interna es fundamental para el éxito, así como porcentaje de ingresos que se invierta en innovación. Además se valida la importancia de haber realizado procesos de innovación en el pasado como condición para un asegurar un proceso de innovación de éxito. Estas conclusiones se derivan de las segmentaciones realizadas.
- Queda en relevancia la importancia de la Financiación Pública en el éxito en la innovación. Además, se comprueba que no es un *false predictor*, ya que, como puede observarse en la siguiente tabla, el porcentaje de empresas que reciben ayudas públicas se duplica en el caso de los éxitos.

	PRE	POST
% Financiación Pública Total Empresas	32%	26%
% Financiación Pública Éxitos	61%	71%

Tabla 13 % Financiación Pública

- Se confirma que las empresas de menor tamaño son las más (Löof y Hesmati, 2004; Wessel, 2012). Estas conclusiones se derivan de las segmentaciones realizadas, y pueden observarse en los siguientes cuadros, para cada uno de los intervalos de análisis.

	Número de Empresas	Tasa de Éxitos
Microempresas	1.504	6,18%
Pequeña Empresa	4.179	4,07%
Mediana Empresa	3.287	2,04%
Gran Empresa	2.615	1,91%

Tabla 14 Tasa de Éxito por Tipo de Empresa (2005-2007)

	Número de Empresas	Tasa de Éxitos
Microempresas	1.601	1,19%
Pequeña Empresa	3.456	0,87%
Mediana Empresa	2.817	0,99%
Gran Empresa	2.075	1,01%

Tabla 15 Tasa de Éxito por Tipo de Empresa (2009-2011)

- Se confirma también la influencia de la edad en el éxito innovador, conclusión que también se deriva de las segmentaciones realizadas. A continuación se muestra la tasa de éxitos para cada uno de los intervalos de madurez definidos y especificados en apartados anteriores.

	Número de Empresas	Tasa de Éxitos	Edad Promedio
1	709	1,64%	3
2	2.661	2,94%	9
3	2.722	0,72%	16
4	2.198	0,49%	24
5	2.516	0,60%	49

Tabla 16 Tasa de Éxito por Edad de Empresa (2005-2007)

	Número de Empresas	Tasa de Éxitos	Edad Promedio
1	61	1,64%	4
2	1.601	2,94%	9
3	2.647	0,72%	16
4	2.636	0,49%	25
5	3.004	0,60%	49

Tabla 17 Tasa de Éxito por Edad de Empresa (2009-2011)

- Así mismo, se valida la influencia del sector de actividad en el perfil de innovación: los sectores de Investigación, Ingeniería y Tecnología son los más innovadores, lo que se justifica por su naturaleza y características, así como por las exigencias y expectativas del mercado al que están destinados.

	Número de Empresas	Tasa de Éxitos
INVESTIGACIÓN	449	12,03%
INGENIERÍA	518	7,92%
TECNOLOGÍA	1.973	4,71%
CONSTRUCCIÓN	466	3,65%
SERVICIOS PRIMARIOS	655	3,51%
FABRICACIÓN	1.705	2,87%

	Número de Empresas	Tasa de Éxitos
MATERIAS PRIMAS	2.347	2,26%
SERVICIOS	1.453	1,79%
TRANSPORTE Y MOTOR	551	1,63%
MANUFACTURAS	619	1,13%
COMERCIO	849	0,94%

Tabla 18 Tasa de Éxito por Sector (2005-2007)

	Número de Empresas	Tasa de Éxitos
INVESTIGACIÓN	235	7,23%
INGENIERÍA	782	1,92%
TECNOLOGÍA	1.800	1,39%
FABRICACIÓN	1.368	0,80%
MATERIAS PRIMAS	2.023	0,69%
COMERCIO	955	0,63%
TRANSPORTE Y MOTOR	498	0,60%
SERVICIOS	1.012	0,40%
CONSTRUCCIÓN	391	0,26%
SERVICIOS PRIMARIOS	429	0,23%
MANUFACTURAS	456	0,22%

Tabla 19 Tasa de Éxito por Sector (2009-2011)

- Queda en relevancia la escasa importancia de la pertenencia a un grupo empresarial o no en cuanto a perfiles innovadores de éxito.

	Número de Empresas	Tasa de Éxitos
Grupo	4.473	3,24%
Sin Grupo	7.112	3,30%

Tabla 20 Tasa de Éxito por Sector (2005-2007)

	Número de Empresas	Tasa de Éxitos
Grupo	4.134	1,14%
Sin Grupo	5.815	0,88%

Tabla 21 Tasa de Éxito por Sector (2009-2011)

6. Próximos Pasos y Principales Limitaciones

Una de las principales limitaciones que se ha encontrado al realizar este trabajo ha sido la falta de variables relacionadas con rentabilidad. Es con el fin de solventar esta carencia, se ha diseñado la variable sintética de *éxito innovador*. No obstante, enriquecería el análisis y las conclusiones obtenidas el disponer de datos procedentes del Balance y la Cuenta de Resultados, de los que derivan las variables relativas a la rentabilidad obtenida en cada ejercicio, como puede ser el ROA o beneficio económico.

Otra limitación que se ha encontrado es la falta de contraste entre la información proporcionada por las empresas como respuesta a la encuesta facilitada por el Instituto Nacional de Estadística y que se recoge en la base de datos utilizada, lo que pone en duda su veracidad y la posible distorsión de la situación real del tejido empresarial español y su grado de innovación. Así mismo, la anonimización de variables realizada con el fin de conservar el secreto estadístico distorsiona los datos obtenidos.

Como Próximos Pasos o líneas de estudio generales, se sugieren tres grandes bloques, siendo conscientes de la amplitud y riqueza de temas derivados de este trabajo. No obstante, como pasos inmediatos se seguirían los siguientes:

- En primer lugar, solventar las principales limitaciones con las que se ha encontrado este trabajo. Esto es:
 - Contraste de la información disponible, mediante la comprobación de la veracidad a través de estudios externos y el contraste con variables ajenas a la empresa o que puedan ser validadas a través de distintas fuentes.
 - Inclusión de variables procedentes del Balance y la Cuenta de Resultados para enriquecer el análisis.
 - Análisis del impacto de la anonimización de variables.
- En segundo caso, se sugiere la profundización en el estudio de la influencia de la Financiación Pública en el éxito innovador:
 - ¿Se trata de un *false predictor* y, como tal, distorsiona el análisis realizado?
 - ¿Tiene efectos aditivos o sustitutivos sobre la inversión en innovación?

- ¿Qué medidas pueden aplicarse para que la Financiación Pública en innovación empresarial sea más efectiva?
- Por último, se sugiere la realización de un análisis detallado del proceso innovador a largo plazo de las empresas por cada segmento definido, así como el diseño de una senda de innovación que permita facilitar el éxito en innovación y la traducción de los esfuerzos innovadores empresariales en resultados para la empresa.

7. Bibliografía

- Alpay, G., M. Bodur, C. Yilmaz y P. Büyükbacı. 2012, “How does innovativeness yield superior firm performance”, *Innovation: Management, policy & practice*, 14 (1), 107-128.
- Aschhoff, B. y Schmidt, T. 2008, “Empirical Evidence on the Success of R&D Cooperation—Happy Together?” *Rev Ind Organ*, 33, 41-62.
- Baptista, R. y P. Swann. 1998, “Do firms in clusters innovate more?”. *Research Policy*, Vol 27, pp. 524-540.
- Barajas, A., E. Huergo y L. Moreno. 2012, “Las empresas españolas que patentan en la Oficina Europea. Características comparadas con las no solicitantes”
- Beard, D.W. y G.G. Dess. 1981, “Corporate-Level Strategy, Business-Level Strategy and Firm Performance”. *The Academy of Management Journal*, Vol.24, No. 4 (Dec., 1981), pp. 663-688.
- Camisón, C. y Monfort-Mir, V. M. 2011, “Measuring innovation in tourism from the Schumpeterian and the dynamic-capabilities perspectives”, *Tourism Management*, 33, 776-789.
- COTEC. 2007, Informes sobre el sistema español de innovación. Innovación en el sector hotelero. Madrid: Fundación COTEC para la Innovación Tecnológica.
- DTI. 2007, Innovation in services. DTI Occasional paper no. 9. London: DTI.
- Flaig, G. y Stadler, M. 1994, “Success Breeds Success. The Dynamics of the Innovation Process”, *Empirical Economics*, 19, 55-68.
- Fundación SEPI. 2015, *La base de datos Panel de Innovación Tecnológica*. Madrid: Instituto Nacional de Estadística.
- Grant, R. M. 2010, *Contemporary Strategy Analysis*. Barcelona: Wiley & Sons, Ltd
- Hagerdoorn, J. 1996, “Innovation and Entrepreneurship: Schumpeter Revisited”, *Industrial and Corporate Change*, 5(3), 883-896.

- Hjalager, A. M. 1994, “Dynamic innovation in the tourism industry”. *Progress in Tourism and Hospitality Management*, 6, 197-224.
- Hjalager, A. M. 2009, “A review of innovation research in tourism”, *Tourism Management*, 31, 1-12.
- Johnson, G., K. Scholes y R. Whittington (2008) *Exploring Corporate Strategy: Text & Cases*. London: Prentice-Hall.
- Leifer, R., G. Colarelli O’Connor y M. Rice. 2001, “Implementing radical innovation in mature firms: The role of hubs”. *Academy of Management Perspectives*, Vol. 15, No. 3.
- Lööf, H. y Heshmati, A. 2005, *The Impact of Public Funding on Private R&D Investment. New Evidence from a Firm Level Innovation Study*. CESIS Electronic Working Paper Series.
- Mariello, A. 2007, “The Five Stages of Successful Innovation”. MIT Sloan Management Review.
- OECD y Eurostat. 2005, Oslo manual: Guidelines for collecting and interpreting innovation data. Paris: OECD.
- Onkelinx, J. y L. Sleuwaegen. 2010, *Internacionalization strategy and performance of small and medium sized enterprises*. Bruselas: National Bank of Belgium
- Orfila-Sintes, F., & J. Mattsson. 2007, “Innovation behavior in the hotel industry”. *Omega*, 37(2), 380–394.
- Parsons, M. y M. B. Rose. 2010, Innovation, Entrepreneurship and Networks. A dance of two questions. En: P. Fernández Pérez y M. B. Rose. *Innovation and Entrepreneurial Networks in Europe*. New York: Routledge International Studies in Business History, 41-60.
- PITEC. 2012, La base de datos Panel de Innovación Tecnológica. Madrid: Instituto Nacional de Estadística.
- Porter, M. E. (1996) “What is Strategy?” *Harvard Business Review*, Nov/Dec 1996

- Rubera, G. y A.H. Kirka. 2012, “Firm Innovativeness and Its Performance Outcomes: A Meta-Analytic Review and Theoretical Integration”, *Journal of Marketing*, 76, 130-147.
- San Román, E., S. Blanco, A. Gil y M. Pérez. 2013, *Is Spain Different? Innovation systems in the long-term development of the tourism sector. The Iberostar case study*.
- Steliana Ciote, C. 2012, “From Schumpeter to neo-Schumpeterianism”, *Economics, Management, and Financial Markets*, 7 (4), 466-472.
- Sundbo, J. (1998). *The Organisation of Innovation in Services*. Copenhagen: Roskilde University Press.
- Sundbo, J.; Orfila-Sintes, F. y Sørensen, F. (2007). “The innovative behaviour of tourism firms –Comparative studies of Denmark and Spain”. *Research Policy*, 36, 88-106.
- Torrent-Sellens, J. y A. Díaz-Chao. 2015, *Las empresas industriales en 2013. Evolución de la crisis e Índice de Competitividad de la Empresa Industrial (ICEMPI)* Madrid: Fundación SEPI.
- Wessel, M. 2012, “Why Big Companies Can’t Innovate”. *Harvard Business Review* (Sep 2012)
- Zoltan J. Acs y David B. Audretsch. 1987, “Innovation, Market Structure, and Firm Size”. *The Review of Economics and Statistics*, Vol. 69, No. 4 (Nov., 1987), pp. 567-574

Anexo I – CNAEs y equivalencias 2005-2007

CNAE_07	CATEGORIA_CNAE	EMPRESAS
ACTIVIDADES POSTALES Y DE CORREO 641	TRANSPORTE Y MOTOR	19
VEHICULOS DE MOTOR 34	TRANSPORTE Y MOTOR	240
ENSAYOS Y ANALISIS TECNICOS 743	INVESTIGACIÓN	146
ACTIVIDADES DE RADIO Y TELEVISION 922	TECNOLOGÍA	38
EXTRACTIVAS 10 a 14	MATERIAS PRIMAS	64
INVESTIGACION Y DESARROLLO 73	INVESTIGACIÓN	303
SERVICIOS DE TELECOMUNICACION 642	TECNOLOGÍA	51
ACTIVIDADES INMOBILIARIAS 70	SERVICIOS	102
OTRAS ACTIVIDADES INFORMATICAS 72 (exc. 722)	TECNOLOGÍA	184
COMERCIO AL POR MAYOR 51	COMERCIO	562
ALIMENTOS Y BEBIDAS 15	MATERIAS PRIMAS	753
PRODUCTOS METALURGICOS NO FERREOS 274, 2753, 2754	MATERIAS PRIMAS	65
TABACO 16	MATERIAS PRIMAS	4
INTERMEDIACION FINANCIERA 65, 66, 67	SERVICIOS	217
JUEGOS Y JUGUETES 365	FABRICACIÓN	18
CUERO Y CALZADO 19	MANUFACTURAS	77
MADERA Y CORCHO 20	MATERIAS PRIMAS	113
INSTRUMENTOS MEDICOS Y DE PRECISION, OPTICOS 33	FABRICACIÓN	248
APARATOS DE RADIO, TV Y COMUNICACION 32 (exc. 321)	TECNOLOGÍA	96
PRODUCTOS FARMACEUTICOS 244	FABRICACIÓN	164
RECICLAJE 37	SERVICIOS PRIMARIOS	42
PAPEL 21	MATERIAS PRIMAS	113
COQUERIAS, REFINO DE PETROLEO 23	MATERIAS PRIMAS	6
OTRO MATERIAL DE TRANSPORTE 35 (exc. 351, 353)	TRANSPORTE Y MOTOR	28
MUEBLES 361	MANUFACTURAS	216
CONSTRUCCION NAVAL 351	INGENIERÍA	46
PRODUCCION Y DISTRIBUCION DE ELECTRICIDAD, GAS Y AGUA 40, 41	SERVICIOS PRIMARIOS	79
ACTIVIDADES ANEXAS AL TRANSPORTE, AGENCIAS DE VIAJE 63	TRANSPORTE Y MOTOR	123
CAUCHO Y MATERIAS PLASTICAS 25	FABRICACIÓN	372
PRODUCTOS METALURGICOS FERREOS 271, 272, 273, 2751, 2752	MATERIAS PRIMAS	101
SERVICIOS TECNICOS DE ARQUITECTURA E INGENIERIA 742	INGENIERÍA	446
VENTA Y REPARACION DE VEHICULOS A MOTOR 50	COMERCIO	85
COMERCIO AL POR MENOR 52	COMERCIO	202
ACTIVIDADES CINEMATOGRAFICAS Y DE VIDEO 921	SERVICIOS	43

CNAE_07	CATEGORIA_CNAE	EMPRESAS
ALQUILER DE MAQUINARIA Y EQUIPO 71	SERVICIOS	44
MAQUINAS DE OFICINA Y EQUIPOS INFORMATICOS 30	TECNOLOGÍA	25
OTRAS ACTIVIDADES SANITARIAS, SOCIALES Y COLECTIVAS 85, 90, 91,92 (exc. 921, 922), 9	SERVICIOS PRIMARIOS	478
CONSTRUCCION 45	CONSTRUCCIÓN	466
AGRICULTURA 01, 02, 05	MATERIAS PRIMAS	171
CONFECCION Y PELETERIA 18	MANUFACTURAS	82
EDUCACION 80 (exc. 8030)	SERVICIOS PRIMARIOS	56
MAQUINARIA Y EQUIPO MECANICO 29	TECNOLOGÍA	810
SOFTWARE 722	TECNOLOGÍA	638
MAQUINARIA Y MATERIAL ELECTRICO 31	FABRICACIÓN	284
QUIMICA (EXCEPTO PRODUCTOS FARMACEUTICOS) 24 (exc. 244)	MATERIAS PRIMAS	604
OTRAS ACTIVIDADES EMPRESARIALES 74 (exc. 742, 743)	SERVICIOS	693
TEXTIL 17	MANUFACTURAS	244
OTRAS MANUFACTURAS 36 (exc. 361, 365)	TECNOLOGÍA	65
EDICION, ARTES GRAFICAS Y REPRODUCCION 22	SERVICIOS	166
PRODUCTOS METALICOS (EXCEPTO MAQUINARIA Y EQUIPO) 28	FABRICACIÓN	619
HOSTELERIA 55	SERVICIOS	188
CONSTRUCCION AERONAUTICA Y ESPACIAL 353	INGENIERÍA	26
AZULEJOS Y BALDOSA CERAMICA 263	MATERIAS PRIMAS	61
TRANSPORTE 60, 61, 62	TRANSPORTE Y MOTOR	141
PRODUCTOS MINERALES NO METALICOS (EXCEPTO AZULEJOS Y BALDOSA CERAMICA) 26 (exc. 263)	MATERIAS PRIMAS	292
COMPONENTES ELECTRONICOS 321	TECNOLOGÍA	66

Anexo II – CNAEs y equivalencias 2009-2011

CNAE	CATEGORIA_CNAE	EMPRESAS
COMERCIO 45, 46, 47	COMERCIO	958
OTROS SERVICIOS 95, 96	COMERCIO	145
CONSTRUCCIÓN 41, 42, 43	CONSTRUCCIÓN	473
FARMACIA 21	FABRICACIÓN	154
OTRAS ACTIVIDADES DE FABRICACIÓN 32	FABRICACIÓN	152
CAUCHO Y PLÁSTICOS 22	FABRICACIÓN	369
MATERIAL Y EQUIPO ELÉCTRICO 27	FABRICACIÓN	284
MANUFACTURAS METÁLICAS 25	FABRICACIÓN	619
OTRAS ACTIVIDADES 69, 70, 71, 73, 74, 75	INGENIERÍA	883
CONSTRUCCIÓN NAVAL 301	INGENIERÍA	26
CONSTRUCCIÓN AERONÁUTICA Y ESPACIAL 303	INGENIERÍA	25
SERVICIOS DE I+D 72	INVESTIGACIÓN	291
CONFECCIÓN 14	MANUFACTURAS	94
MUEBLES 31	MANUFACTURAS	201
CUERO Y CALZADO 15	MANUFACTURAS	63
TEXTIL 13	MANUFACTURAS	207
CARTÓN Y PAPEL 17	MATERIAS PRIMAS	106
METALURGIA 24	MATERIAS PRIMAS	160
QUÍMICA 20	MATERIAS PRIMAS	570
MADERA Y CORCHO 16	MATERIAS PRIMAS	104
ALIMENTACIÓN, BEBIDAS Y TABACO 10, 11, 12	MATERIAS PRIMAS	752
AGRICULTURA, GANADERÍA, SILVICULTURA Y PESCA 01, 02, 03	MATERIAS PRIMAS	146
PRODUCTOS MINERALES NO METÁLICOS DIVERSOS 23	MATERIAS PRIMAS	328
INDUSTRIAS DEL PETRÓLEO 19	MATERIAS PRIMAS	3
INDUSTRIAS EXTRACTIVAS 05, 06, 07, 08, 09	MATERIAS PRIMAS	58
HOSTELERÍA 55, 56	SERVICIOS	191
ACTIVIDADES ADMINISTRATIVAS Y SERVICIOS AUXILIARES 77, 78, 79, 80, 81, 82	SERVICIOS	493
ACTIVIDADES FINANCIERAS Y DE SEGUROS 64, 65, 66	SERVICIOS	246
ACTIVIDADES INMOBILIARIAS 68	SERVICIOS	96
ACTIVIDADES ARTÍSTICAS, RECREATIVAS Y DE ENTRETENIMIENTO 90, 91, 92, 93	SERVICIOS	55
ARTES GRÁFICAS Y REPRODUCCIÓN 18	SERVICIOS	89
EDUCACIÓN 85 (EXC. 854)	SERVICIOS PRIMARIOS	61
ENERGÍA Y AGUA 35, 36	SERVICIOS PRIMARIOS	81
ACTIVIDADES SANITARIAS Y DE SERVICIOS SOCIALES 86, 87, 88	SERVICIOS PRIMARIOS	245
SANEAMIENTO, GESTIÓN DE RESIDUOS Y DESCONTAMINACIÓN 37, 38, 39	SERVICIOS PRIMARIOS	104
REPARACIÓN E INSTALACIÓN DE MAQUINARIA Y EQUIPO 33	TECNOLOGÍA	111
PRODUCTOS INFORMÁTICOS, ELECTRÓNICOS Y ÓPTICOS 26	TECNOLOGÍA	318

CNAE	CATEGORIA_CNAE	EMPRESAS
OTRA MAQUINARIA Y EQUIPO 28	TECNOLOGÍA	720
PROGRAMACIÓN, CONSULTORÍA Y OTRAS ACTIVIDADES INFORM62	TECNOLOGÍA	671
TELECOMUNICACIONES 61	TECNOLOGÍA	62
OTROS SERVICIOS DE INFORMACIÓN Y COMUNICACIONES 58, 59, 60, 63	TECNOLOGÍA	293
VEHÍCULOS DE MOTOR 29	TRANSPORTE Y MOTOR	272
OTRO EQUIPO DE TRANSPORTE 30 (EXC. 301, 303)	TRANSPORTE Y MOTOR	34
TRANSPORTES Y ALMACENAMIENTO 49, 50, 51, 52, 53	TRANSPORTE Y MOTOR	249

Anexo III – Índice concentración 2005-2007

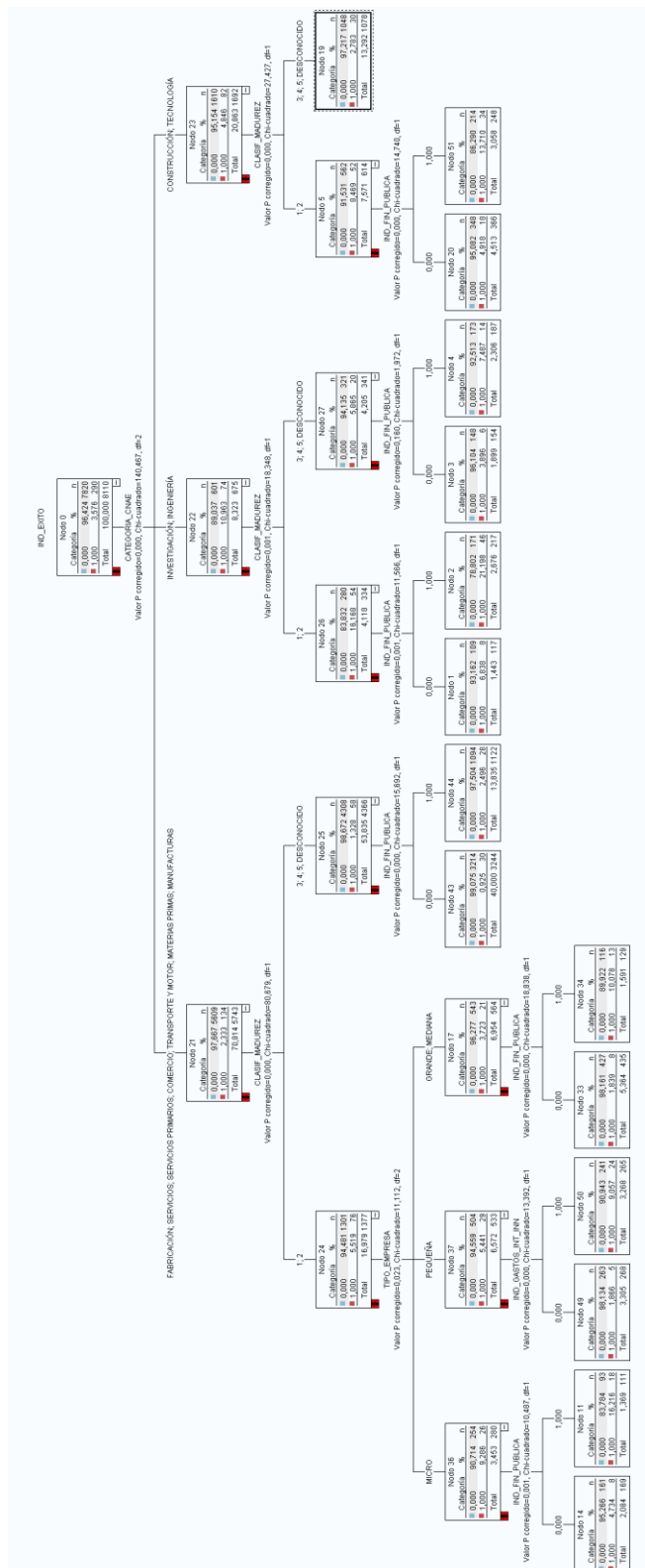
Agrupación de CNAE	Índice de Concentración	Número de Empresas	Clasificación
TABACO 16	2.489	14	1
COQUERIAS, REFINO DE PETROLEO 23	2.407	19	1
JUEGOS Y JUGUETES 365	1.498	54	1
ACTIVIDADES POSTALES Y DE CORREO 641	1.370	75	1
SERVICIOS DE TELECOMUNICACION 642	1.263	79	1
CONSTRUCCION AERONAUTICA Y ESPACIAL 353	1.658	57	1
MAQUINAS DE OFICINA Y EQUIPOS INFORMATICOS 30	1.407	177	1
ACTIVIDADES DE RADIO Y TELEVISION 922	1.183	118	1
CONSTRUCCION NAVAL 351	1.084	136	1
OTRO MATERIAL DE TRANSPORTE 35 (exc. 351, 353)	1.070	92	1
APARATOS DE RADIO, TV Y COMUNICACION 32 (exc. 321)	934	261	2
VENTA Y REPARACION DE VEHICULOS A MOTOR 50	929	530	2
OTRAS ACTIVIDADES INFORMATICAS 72 (exc. 722)	921	254	2
PRODUCCION Y DISTRIBUCION DE ELECTRICIDAD, GAS Y AGUA 40, 41	901	135	2
ALQUILER DE MAQUINARIA Y EQUIPO 71	781	201	2
PRODUCTOS METALURGICOS NO FERREOS 274, 2753, 2754	751	132	2
ACTIVIDADES CINEMATOGRAFICAS Y DE VIDEO 921	749	270	2
CONFECCION Y PELETERIA 18	712	166	2
EDUCACION 80 (exc. 8030)	648	119	2
ACTIVIDADES INMOBILIARIAS 70	565	410	2
COMERCIO AL POR MENOR 52	563	216	2
RECICLAJE 37	997	316	2
VEHICULOS DE MOTOR 34	700	270	2
TRANSPORTE 60, 61, 62	630	753	2
COMPONENTES ELECTRONICOS 321	649	668	2
MADERA Y CORCHO 20	512	757	2
INSTRUMENTOS MEDICOS Y DE PRECISION, OPTICOS 33	317	591	3
AGRICULTURA 01, 02, 05	312	492	3
CONSTRUCCION 45	279	1.138	4
COMERCIO AL POR MAYOR 51	179	383	3
TEXTIL 17	114	190	3
ALIMENTOS Y BEBIDAS 15	79	875	3
PRODUCTOS METALICOS (EXCEPTO MAQUINARIA Y EQUIPO) 28	73	203	3
AZULEJOS Y BALDOSA CERAMICA 263	344	188	3

ENSAYOS Y ANALISIS TECNICOS 743	344	454	3
INVESTIGACION Y DESARROLLO 73	331	913	3
INTERMEDIACION FINANCIERA 65, 66, 67	221	911	3
HOSTELERIA 55	192	1.853	4
EDICION, ARTES GRAFICAS Y REPRODUCCION 22	268	1.407	4
CAUCHO Y MATERIAS PLASTICAS 25	278	2.516	4
MAQUINARIA Y EQUIPO MECANICO 29	138	308	3
SERVICIOS TECNICOS DE ARQUITECTURA E INGENIERIA 742	320	677	3
MUEBLES 361	217	2.133	4
OTRAS ACTIVIDADES SANITARIAS, SOCIALES Y COLECTIVAS 85, 90, 91,92 (exc. 921, 922), 9	260	661	3
SOFTWARE 722	418	1.383	4
PRODUCTOS MINERALES NO METALICOS (EXCEPTO AZULEJOS Y BALDOSA CERAMICA) 26 (exc. 263)	186	237	3
OTRAS ACTIVIDADES EMPRESARIALES 74 (exc. 742, 743)	132	350	3
QUIMICA (EXCEPTO PRODUCTOS FARMACEUTICOS) 24 (exc. 244)	109	509	3
CUERO Y CALZADO 19	416	1.771	4
ACTIVIDADES ANEXAS AL TRANSPORTE, AGENCIAS DE VIAJE 63	403	776	3
PRODUCTOS METALURGICOS FERREOS 271, 272, 273, 2751, 2752	390	2.320	4
PAPEL 21	350	1.881	4
EXTRACTIVAS 10 a 14	534	346	2
PRODUCTOS FARMACEUTICOS 244	243	2.006	4
MAQUINARIA Y MATERIAL ELECTRICO 31	254	1.475	4
OTRAS MANUFACTURAS 36 (exc. 361, 365)	462	500	3

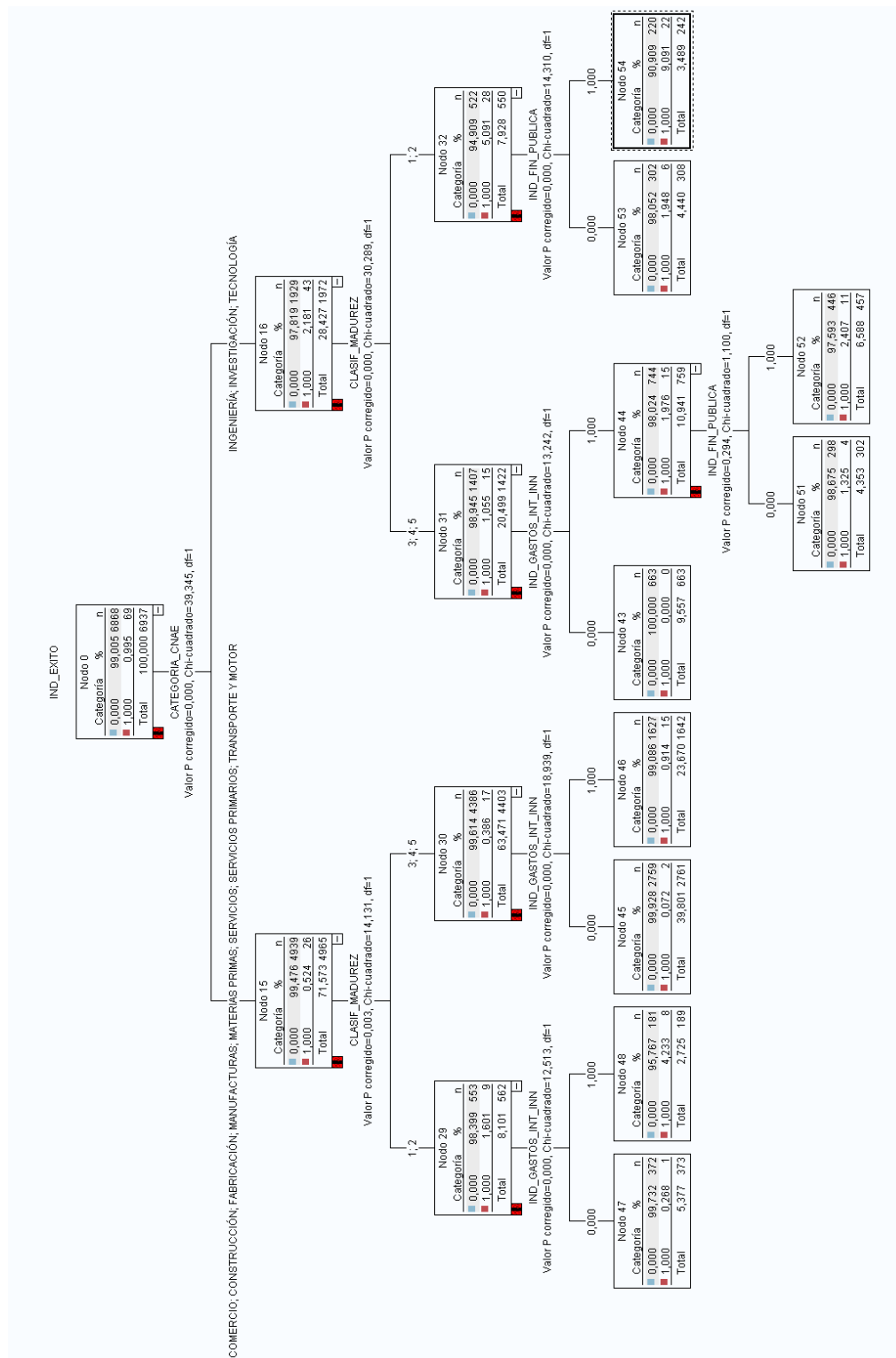
Anexo IV – Índice concentración 2009-2011

Agrupación de CNAE	Índice de Concentración	Número de Empresas	Clasificación
INDUSTRIAS DEL PETRÓLEO 19	3.375	7	1
CONSTRUCCIÓN AERONÁUTICA Y ESPACIAL 303	1.582	60	1
TELECOMUNICACIONES 61	1.435	163	1
OTRO EQUIPO DE TRANSPORTE 30 (exc. 301, 303)	1.260	86	1
CONSTRUCCIÓN NAVAL 301	1.227	68	1
ACTIVIDADES ARTÍSTICAS, RECREATIVAS Y DE ENTRETENIMIENT 90, 91, 92, 93	888	148	2
ENERGÍA Y AGUA 35, 36	845	224	2
INDUSTRIAS EXTRACTIVAS 05, 06, 07, 08, 09	837	164	2
ARTES GRÁFICAS Y REPRODUCCIÓN 18	778	241	2
SANEAMIENTO, GESTIÓN DE RESIDUOS Y DESCONTAMINACIÓN 37, 38, 39	750	264	2
CUERO Y CALZADO 15	703	169	2
CONFECCIÓN 14	691	254	2
VEHÍCULOS DE MOTOR 29	658	741	2
ACTIVIDADES INMOBILIARIAS 68	616	210	2
EDUCACIÓN 85 (exc. 854)	538	160	2
MADERA Y CORCHO 16	525	279	2
OTROS SERVICIOS DE INFORMACIÓN Y COMUNICACIONES 58, 59, 60, 63	511	754	2
OTROS SERVICIOS 95, 96	447	338	3
TRANSPORTES Y ALMACENAMIENTO 49, 50, 51, 52, 53	443	704	3
CARTÓN Y PAPEL 17	432	297	3
MATERIAL Y EQUIPO ELÉCTRICO 27	412	766	3
PROGRAMACIÓN, CONSULTORÍA Y OTRAS ACTIVIDADES INFORM62	408	1.749	4
OTRAS ACTIVIDADES DE FABRICACIÓN 32	378	397	3
REPARACIÓN E INSTALACIÓN DE MAQUINARIA Y EQUIPO 33	374	259	3
AGRICULTURA, GANADERÍA, SILVICULTURA Y PESCA 01, 02, 03	363	400	3
METALURGIA 24	314	450	3
FARMACIA 21	309	432	3
ACTIVIDADES SANITARIAS Y DE SERVICIOS SOCIALES 86, 87, 88	306	692	3
CAUCHO Y PLÁSTICOS 22	297	1.026	4
SERVICIOS DE I+D 72	297	771	3
PRODUCTOS INFORMÁTICOS, ELECTRÓNICOS Y ÓPTICOS 26	289	805	3

Agrupación de CNAE	Índice de Concentración	Número de Empresas	Clasificación
ACTIVIDADES FINANCIERAS Y DE SEGUROS 64, 65, 66	269	656	3
CONSTRUCCIÓN 41, 42, 43	259	1.260	4
MUEBLES 31	238	520	3
COMERCIO 45, 46, 47	234	2.552	4
HOSTELERÍA 55, 56	208	551	3
ACTIVIDADES ADMINISTRATIVAS Y SERVICIOS AUXILIARES 77, 78, 79, 80, 81, 82	207	1.360	4
OTRAS ACTIVIDADES 69, 70, 71, 73, 74, 75	194	2.310	4
PRODUCTOS MINERALES NO METÁLICOS DIVERSOS 23	156	917	3
TEXTIL 13	153	563	3
OTRA MAQUINARIA Y EQUIPO 28	131	1.965	4
MANUFACTURAS METÁLICAS 25	129	1.670	4
QUÍMICA 20	127	1.607	4
ALIMENTACIÓN, BEBIDAS Y TABACO 10, 11, 12	86	2.144	4



Anexo VI – Árbol de Decisión 2009-2011



- Script de carga y normalización de los datos



20151023_PREPARACION DATOS_VF.sql

- Script de creación de tablón 2005-2007



20151031_CREACION TABLON ANALISIS PRE_V2.sql

- Script de creación de tablón 2009-2011



20151101_CREACION TABLON ANALISIS POST_V1.sql

- Script de generación y análisis de segmentos



20151110_ANÁLISIS SEGMENTOS_V2.sql